



Grado en Ingeniería de Sistemas Audiovisuales
2016-2017

Trabajo Fin de Grado

“Desarrollo de una app para personas invidentes de reconocimiento de emociones a partir de expresiones faciales”

Nahuel Ignacio Pardo Abbiatti

Tutor:
David Griol Barres

Septiembre 2017. Leganés, Madrid, España.

Agradecimientos:

-En primer lugar quiero agradecerles a mis padres. A medida que he ido creciendo he asimilado las apuestas y sacrificios que habéis realizado en vuestras vidas en favor de mis hermanos y yo. Gracias también por vuestra confianza y apoyo a lo largo de esta carrera de fondo.

A mis hermanos por su comprensión, su confianza en mis facultades y por haberme influenciado en el gusto por el tema de la informática desde pequeño.

Al resto de mi familia, la que vive en México y Argentina. Por vuestro perenne interés sobre mi devenir.

A mis amigos, en especial a Eva, Jara, Víctor, Álvaro, Gonzalo, Irene, Javi y Charly. Con ellos he ido construyendo mi personalidad, mis inquietudes, y he aprendido muchísimo gracias a todos y cada uno de ellos. Os admiro y os quiero.

A mi tutor David por la paciencia mostrada conmigo y a todos los profesores con los que he tratado en la universidad, de todos ellos he sacado algo positivo que me han permitido llegar hasta aquí.

Y quiero finalizar estos agradecimientos nombrando a la persona más generosa y con el corazón más grande con quien he tenido la suerte de topar en mi temprana vida. Sin él nada de esto habría sido posible y nunca podré devolverle toda la ayuda que me ha prestado. Este proyecto y esta carrera va dedicada a él.

Gracias Enrique.

Resumen:

Gracias a los avances en diversos campos de la ciencia e ingeniería (biomedicina, mecánica, informática...) se ha logrado que personas que sufren distintas y complejas discapacidades obtengan una notable mejora en su calidad de vida. Prótesis, audífonos, órganos artificiales, sensores, son solo algunos ejemplos que nos ha dado la tecnología para ayudar a personas con discapacidades muy diversas.

En este proyecto nos centraremos en la discapacidad que sufren más de 55.000 personas en nuestro país, la ceguera. Analizaremos el estado del arte sobre la tecnología destinada a este tipo de discapacidad, a su naturaleza y previsiones en el número de afectados. Nos centraremos en el campo de la telefonía móvil. Estos dispositivos son cada vez más sofisticados y disponibles a más parte de la población, eso que lo convierte en un elemento crucial para facilitar la vida de cualquier discapacitado, no sólo invidentes.

El presente proyecto bebe de las tecnologías potenciadas en las últimas décadas relacionadas con el análisis de sentimientos y reconocimiento de emociones. De todas las subcategorías existentes nos focalizaremos en el análisis de expresiones faciales y eventos gestuales. Realizaremos una aplicación móvil utilizando bibliotecas de Affectiva para crear un traductor de expresiones, apariencia física y eventos faciales a voz. Esto permitirá a una persona invidente obtener información ambiental y visual de su interlocutor, lo cual, debido a esta discapacidad, le resultaría imposible de adquirir de otra forma.

Las aplicaciones que se pueden sustraer de esta herramienta son muy variadas pero todas ellas con un punto en común: obtener información de parte de la comunicación corporal.

Esta forma de comunicación, por definición, está desligada para las personas con esta discapacidad (salvo eventos que involucren contacto físico). Es por ello que se abre un abanico de posibilidades muy disperso. A modo de ejemplo: en el ámbito familiar sería posible obtener información de los gestos faciales que realiza nuestro interlocutor al reaccionar sobre un evento. En un ámbito laboral sería posible obtener información de forma **independiente** de la apariencia de nuestro interlocutor lo que conllevaría, por ejemplo, a tomar medidas en la manera de expresarnos de ahí en adelante.

En definitiva las ventajas rondarán a toda aquella información que las personas sin discapacidades visuales severas podemos obtener de nuestros interlocutores sin mediar palabra, simplemente observando la apariencia y gesticulación facial.

Dado que esta discapacidad es universal y no distingue entre nacionalidades ni lenguajes, la aplicación se ha traducido a ocho idiomas.

Palabras clave:

Reconocimiento de emociones, Android, Discapacidad visual, Texto a voz.

Abstract:

Thanks to the progress on many fields of science and engineering (such as biomedical, mechanics, computing...) it has been possible to achieve that people who suffer from many kinds and complex disabilities get a significant improvement on their quality of life. Prothesis, hearing aids, artificial organs, sensors... they are some few examples of what technology have given us for helping people with different disabilities.

In this project we are going to study the disability suffered from more than 55.000 people in our country: the blindness. We are going to analyze the state of the art of technology used on this disability, its nature and predictions on the number of patients. We are going to focus on the mobile phone's field. These devices are more and more sophisticated and available for the population, that makes it an important factor for making easier the life of any disabled.

The current project feeds from the lately strengthened technologies used on sentiment analysis and emotion recognition. Among all existing subcategories, we are going to focus on facial expressions analysis and gestural events. We will build an app using resources from Affectiva for creating a translator of expressions, physical appearance and facial gestural events to voice. This will allow to a blind person to get ambiantal and visual information of his interlocutor, it would be impossible to get otherwise due to his disability.

The applications that anyone can obtain from this tool are very diverse, but all of them have a point in common: to obtain information from part of the body communication.

This way of communicate, by definition, is indifferent to blind people (except those that involve physical contact). It is therefore a wide range of possibilities. For example: In a familiar environment it would be possible to extract information from facial gestures done by our interlocutor when he reacts over some event. In a work environment, it would be possible to get information **independently** from, for example, the appearance of our client, which can leads to take different strategies in our speech form then to on.

Definitely, the advantages will go round the idea of get every kind of information that a person without any severe visual disability can get from his interlocutor without verbal communication, only watching to his appearance and facial gesture.

Given that this disability is universal and it does not distinguish between languages, the app has been translated into eight languages.

Key words:

Emotion recognition, Android, Visual impairment, Text to speech.

Tabla de contenidos

Capítulo 1: INTRODUCCIÓN	12
1.1 Motivación y objetivos	12
1.2 Recursos utilizados	18
1.3 Fases de desarrollo	19
1.4 Estructura de la memoria	20
Capítulo 2: ESTADO DEL ARTE	22
2.1 Sistema operativo: Android	22
2.2 Texto a Voz	25
2.3 Análisis de sentimientos y reconocedor de emociones	25
2.4 Afectiva	28
2.5 Panorama de aplicaciones para invidentes	33
2.6 Marco Regulatorio	35
Capítulo 3: CICLO DE VIDA DEL PROYECTO	37
3.1 Características de la aplicación	37
3.2 Ciclo de vida del proyecto	38
Capítulo 4: DESCRIPCIÓN DEL SISTEMA	40
4.1 Descripción de la aplicación	40
4.2 Especificación de requerimientos	41
4.2.1 Requerimientos técnicos	41
4.2.1.1 Dispositivo	41
4.2.1.2 Google TalkBack	42
4.2.1.3 SO Android	43
4.2.2 Requerimientos ambientales	43
4.2.2.1 Luz natural	43
4.2.2.2 Luz artificial	44
Capítulo 5: FUNCIONAMIENTO DEL SISTEMA	46
5.1 Selección de idioma	46
5.2 Estados	47
5.2.1 Localización	47
5.2.2 Descripción de apariencia	48
5.2.3 Descriptor de emociones	50
5.2.4 Eventos	51
5.3 Interfaz	52
Capítulo 6: EVALUACIÓN DEL SISTEMA	56
6.1 Resultados de la evaluación	56
6.2 Conclusiones de la evaluación	59

Capítulo 7: CONCLUSIONES Y TRABAJO FUTURO	60
7.1 Conclusiones	60
7.2 Trabajo futuro	61
Capítulo 8: GESTIÓN DEL PROYECTO	63
8.1 Planificación temporal	63
8.2 Presupuesto	64
8.3 Impacto socio-económico	66
ANEXO A: Fases de desarrollo	67
ANEXO B: English content	68
BIBLIOGRAFÍA	77

Índice de figuras:

Figura 1.1 Discapacitados visuales (en millares) según edad.

Figura 1.2 Incapacitados (en millares) para percibir cualquier imagen según edad.

Figura 1.3 Tasa de afiliados a la ONCE según edad.

Figura 1.4 Tasa de afiliados a la ONCE por grado de discapacidad.

Figura 1.5 Proporción de discapacitados visuales en base a la población mundial (millones).

Figura 1.6 Proporción de ceguera y baja visión (millones).

Figura 1.7 Distribución territorial de discapacidad visual.

Figura 2.1 Cuota de mercado de sistemas operativos móviles en smartphone.

Figura 2.2 Previsualización de algoritmos de rastreo de expresiones faciales.

Figura 2.3 Dispositivo para la asistencia de lectura FingerReader.

Figura 2.4 Captura del vídeo promocional de KNFB Reader.

Figura 4.1 Captura de Ajustes/Accesibilidad, donde se podrá activar TalkBack.

Figura 4.2 Captura de los resultados ante un escenario con luz natural incidiendo frontalmente.

Figura 4.3 Captura de los resultados ante un escenario con luz natural incidiendo de espaldas.

Figura 4.4 Captura de los resultados ante un escenario con luz artificial incidiendo frontalmente.

Figura 4.5 Captura de los resultados ante un escenario con luz artificial incidiendo lateralmente.

Figura 5.1 Cambiar el lenguaje del SO y de la Síntesis de voz.

Figura 5.2 Disposición de valores de idiomas en Android Studio.

Figura 5.3 Código relativo a la no-localización de imagen facial.

Figura 8.1 Diagrama de Gantt del proyecto.

Figura 8.2 Desglose de las tareas del proyecto.

Índice de tablas:

Tabla 1.1 Desglose de discapacitados visuales (en millares).

Tabla 2.1 Reparto internacional de versiones Android en 2017.

Tabla 2.2 Arquitectura de software de Android.

Tabla 2.3 Etnias del identificador de etnia de Affective.

Tabla 2.4 Emociones del identificador de emociones de Affective.

Tabla 2.5 Emojis del identificador de emojis de Affective.

Tabla 4.1 Tabla de gestos y acciones de Google TalkBack.

Tabla 5.1 Iconos designados para emociones y eventos.

Tabla 8.1: Amortización de las herramientas utilizadas.

Tabla 8.2: Costes totales del proyecto.

Capítulo 1

INTRODUCCIÓN

En este capítulo se detallarán las circunstancias que me llevaron a enfocar y resolver la problemática del proyecto. Se explicará el contexto social y tecnológico relacionado con la ceguera, aportando estadísticas nacionales e internacionales sobre la discapacidad y ejemplos de aplicaciones tecnológicas destinadas a ella.

Para finalizar se presentarán las fases de desarrollo del proceso de elaboración del proyecto, explicando detenidamente cada una de ellas; y la estructura de la presente memoria que contendrá una breve descripción de cada capítulo.

1.1 Motivación y objetivos

-Junto con el avance exponencial en el campo de la computación, tratamiento de información y conexiones telemáticas de las últimas décadas, nuestra vida cotidiana ha cambiado notablemente. Se han introducido en nuestro día a día nuevas formas de comunicación, más facilidades para comunicarnos a miles de kilómetros de distancia, instantaneidad de nuestras ideas... Hemos sido testigos de cómo, en el periodo de tiempo de nuestra existencia, hemos pasado de mandar cartas a otro continente sin saber si quiera si llegarán, a poder compartir un desayuno con un familiar por videoconferencia, gratuitamente, utilizando nuestro teléfono móvil.

Sin embargo, ¿qué hay de la persona que, antes del boom tecnológico, tenía dificultades para comunicar ideas más básicas que todo eso? Una sonrisa, un beso, un tono de voz, un aspaviento, son solo algunos ejemplos de mensajes que son triviales para la mayoría de nosotros y que utilizamos constantemente en nuestro día a día, pero que para mucha gente con discapacidad es un deseo el experimentarlo.

¿Habrán sido capaces estos avances tecnológicos de solventar esto?

discapacitado, da:

Dicho de una persona: Que padece una disminución física, sensorial o psíquica que la incapacita total o parcialmente para el trabajo o para otras tareas ordinarias de la vida.

De todas las discapacidades que afectan a las comunicaciones de ideas básicas nos vamos a centrar en la **invidencia**. El motivo es sencillo. Todas estas nuevas herramientas de comunicación expresen al máximo la información visual. ¿Qué mejor manera de saber si estos avances ayudan a resolver la problemática que viendo si es capaz de hacerlo en un campo donde la información visual no se aplica?

Estudiaremos entonces el contexto de esta discapacidad a nivel nacional y mundial para obtener pistas que nos puedan servir a la hora de afrontar un diseño y desarrollo de una aplicación móvil.

- *Ámbito nacional*

Conoceremos a continuación datos que nos permitan discernir un número aproximado de personas a las cuales les afectaría este proyecto. Para ello se utilizarán dos fuentes de información distintas. La primera de ellas será del Instituto Nacional de Estadística (INE) y se basará en un estudio realizado en 2008. La segunda provendrá de la Organización Nacional de Ciegos Españoles (ONCE) y nos ofrecerá datos relativos a sus afiliados de este último año.

i) Datos del Instituto Nacional de Estadística [1]

De sus datos podemos obtener el número de afectados por diversas discapacidades, pero centrándonos en las relativas a la visión se disciernen tres agrupaciones: las relacionadas con la imposibilidad de realizar tareas visuales de detalle, de conjunto y finalmente percibir cualquier cualquier tipo de imagen.

Cabe resaltar la distribución en edad de estos datos. Esto nos permitirá obtener cierta información del usuario final que una potencial aplicación. Distribuyendo entonces de esta manera la información obtenemos las siguientes gráficas:

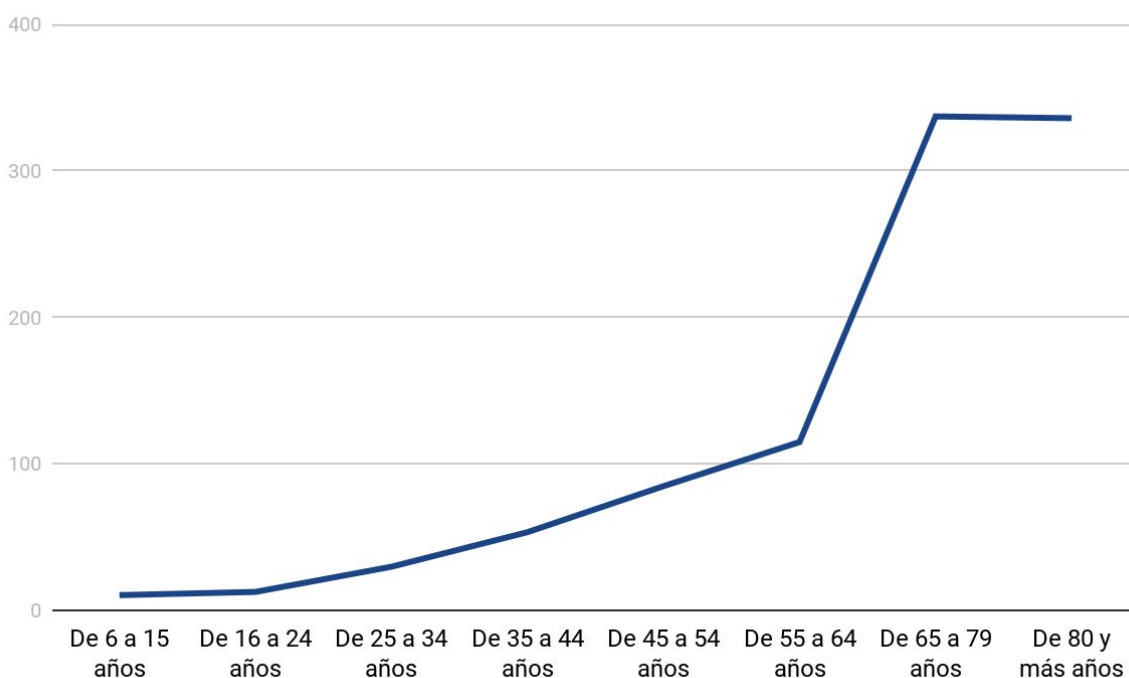


Figura 1.1 Discapacitados visuales (en millares) según edad

En la *figura 1.1* se muestran los afectados de cualquier discapacidad visual, esto agrupa los tres tipos mencionados anteriormente. Se comprueba un brusco ascenso a partir de los 65 años, de tal modo que el 68% del total está comprendido en las edades superiores a los 65 años.

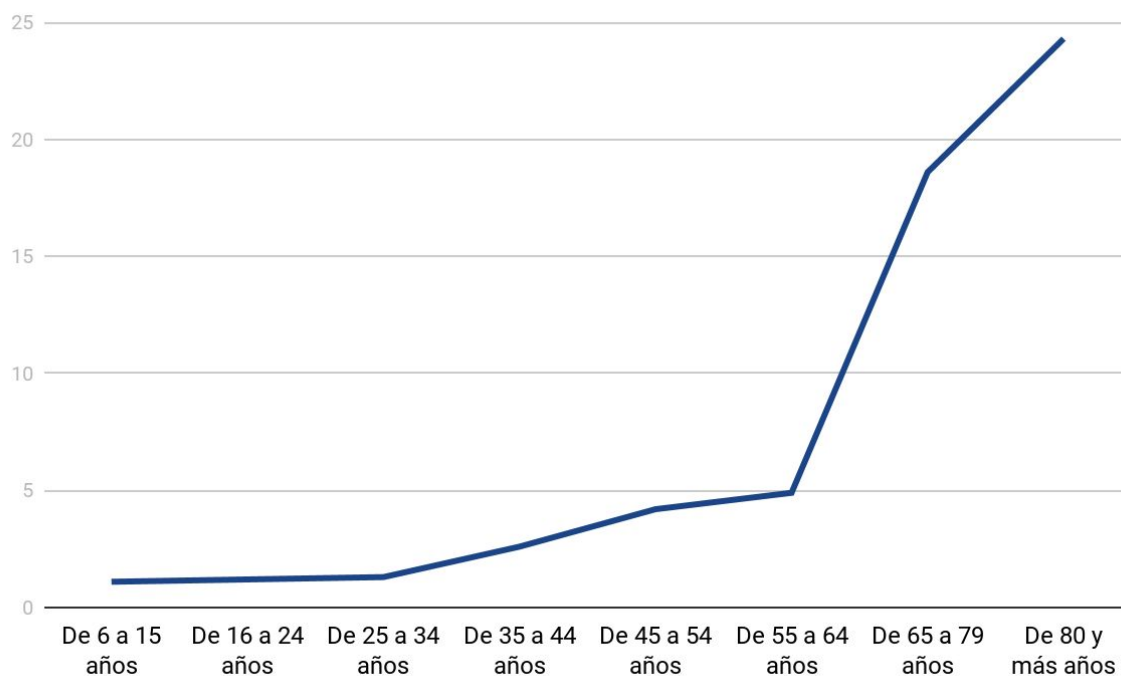


Figura 1.2 Incapacitados (en millares) para percibir cualquier imagen según edad

En la *figura 1.2* se muestra el número de discapacitados de exclusivamente aquellos que son incapaces de percibir cualquier imagen. Una vez más se comprueba esa tendencia acentuada al alza a partir de los 65 años, siendo del 73% del total el rango comprendido desde esta edad en adelante.

En la *tabla 1.1* se muestra el número absoluto de afectados según edad y agrupación de discapacidad:

	Total	De 6 a 64 años	De 65 a 79 años	De 80 y más años
	Ambos sexos	Ambos sexos	Ambos sexos	Ambos sexos
Visión	979,2	306,0	337,2	336,0
Percibir cualquier imagen	58,3	15,4	18,6	24,3
Tareas visuales de detalle	673,6	188,3	241,2	244,0
Tareas visuales de conjunto	662,1	200,4	223,1	238,6
Otros problemas de visión	357,4	127,6	117,7	112,2

Tabla 1.1 Desglose de discapacitados visuales (en millares)

La tasa en base a la población española de aquellas personas con incapacidad de percibir cualquier imagen es del 0,13% y del resto de agrupaciones juntas (baja visión) es del 2.01%.

ii) Datos de los afiliados a la ONCE

La ONCE posee en junio de 2017 un total de 72.097 afiliados [2].

En la *figura 1.3* se distribuyen los afiliados según la edad:

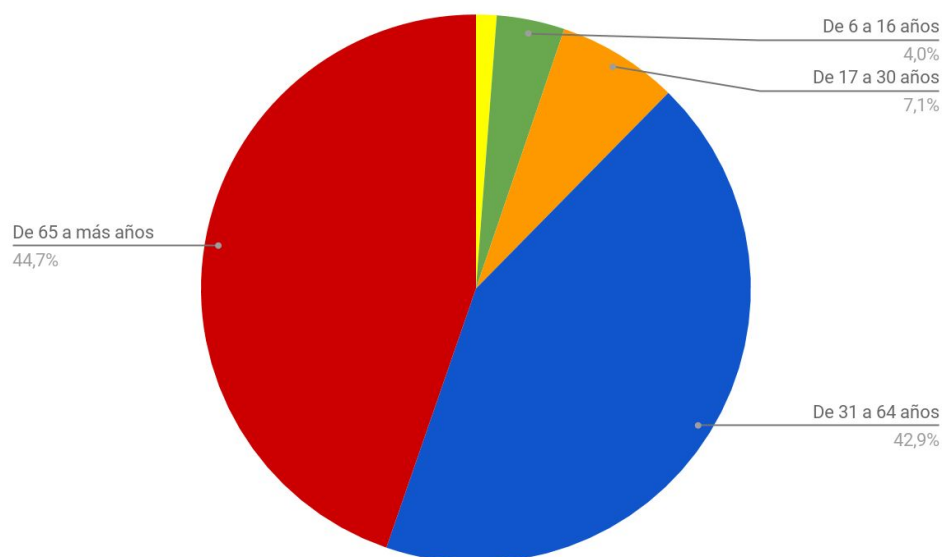


Figura 1.3 Tasa de afiliados a la ONCE según edad

Se comprueba que la mayor parte de afiliados es mayor de los 65 años, siendo la segunda las edades comprendidas entre los 31 y 64 años.

Y en la figura 1.4 podemos distribuirlo según grado de discapacidad:

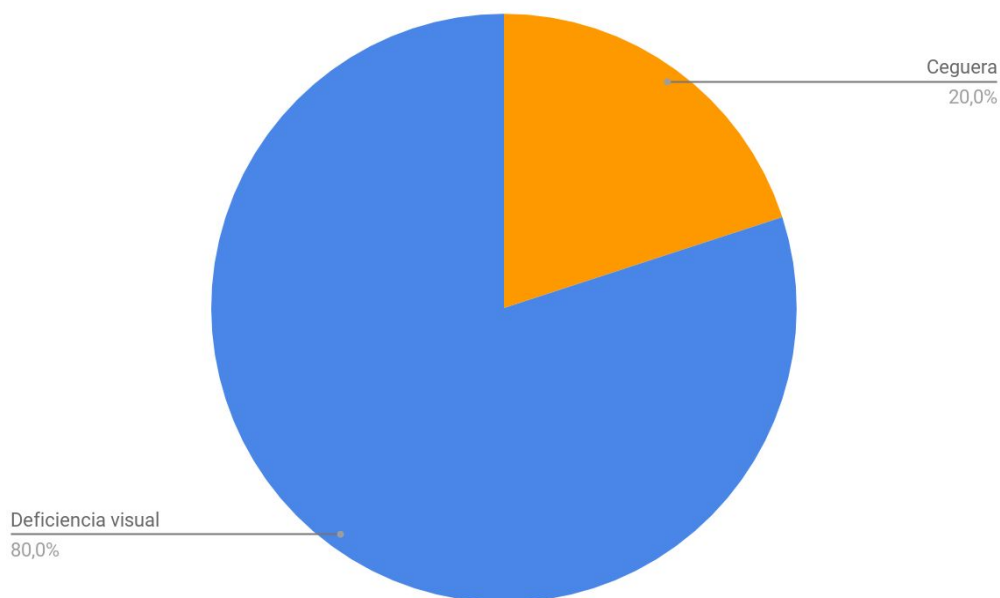


Figura 1.4 Tasa de afiliados a la ONCE por grado de discapacidad.

Definiendo ‘ceguera’ como ausencia de visión o sólo percepción de luz, y ‘deficiencia visual’ como capacidad de mantener un resto visual funcional para la vida diaria.

En definitiva se obtiene que 14.419 afiliados a la ONCE sufre ceguera total.

- *Ámbito internacional*

Bajo este contexto la Organización Mundial de la Salud (OMS) está llevando a cabo un plan de acción mundial para 2014-2019 sobre salud ocular universal [3][4]. En él se nos ofrece los siguientes datos relativos a 2010:

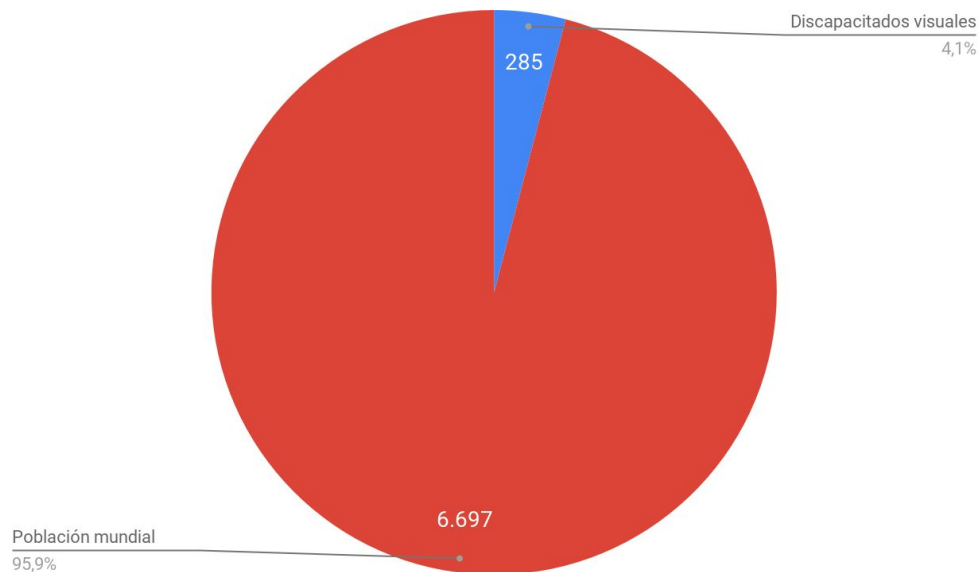


Figura 1.5 Proporción de discapacitados visuales en base a la población mundial (millones).

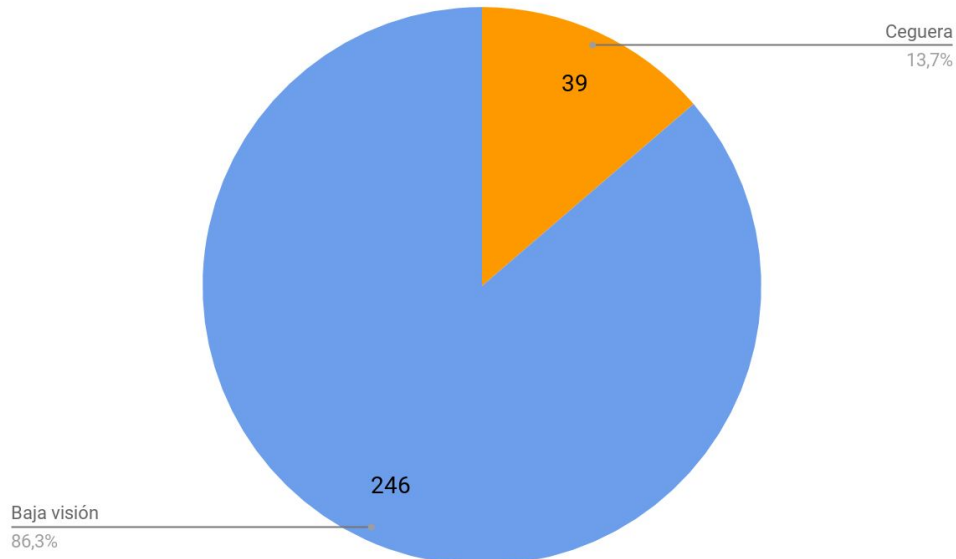


Figura 1.6 Proporción de ceguera y baja visión (millones).

En la *figura 1.5* podemos ver que un 4.1% de la población mundial posee algún tipo de discapacidad visual, lo que equivale a 285 millones de personas.

En la *figura 1.6* desglosamos esas 285 millones de personas según su grado de discapacidad, dividiéndolos en discapacitados con baja visión y discapacitados con ceguera. Un total de 246 millones de personas poseen problemas de baja visión, y la cantidad restante, 39 millones, tendrán una pérdida completa de la visión.

Hay que destacar también cómo están distribuidos estos datos en el territorio mundial:

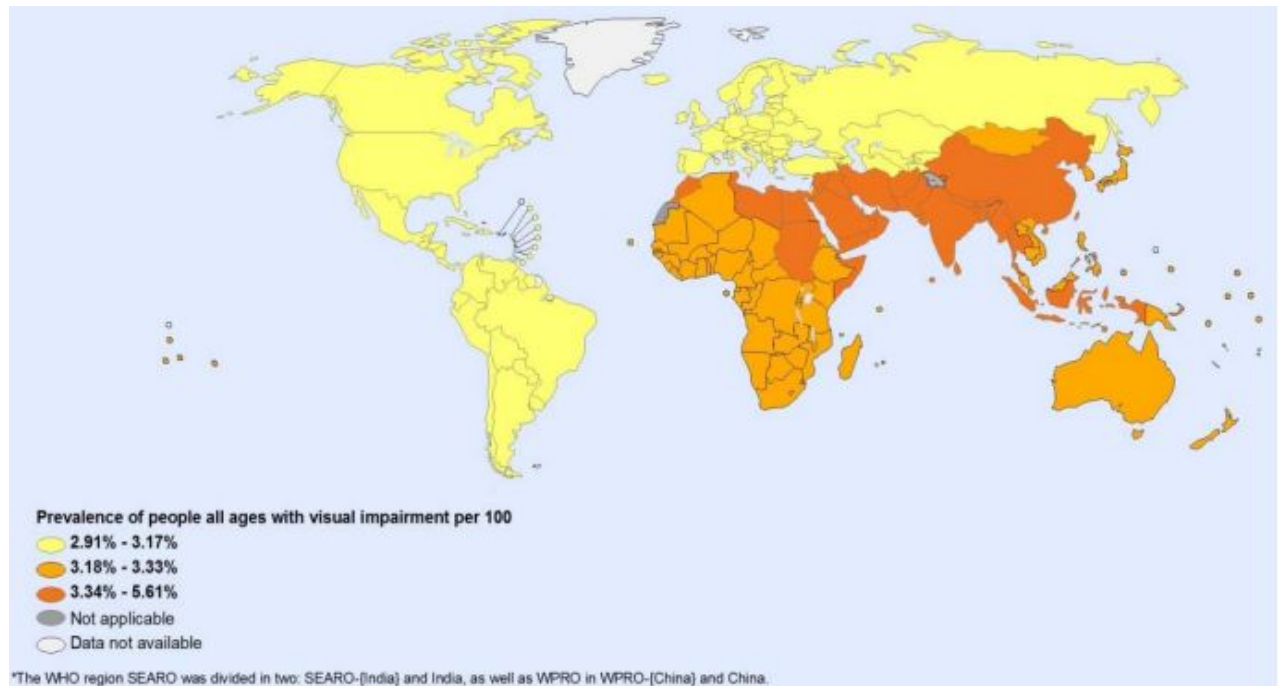


Figura 1.7 Distribución territorial de discapacidad visual.

De la *figura 1.7* podemos sustraer la conclusión de que los países más afectados provienen del continente africano, Arabia, sur y sudeste asiático y Oceanía.

Eso nos da una pista sobre qué lenguajes utilizar en caso de querer abrir el abanico de posibles usuarios de la aplicación que se vaya a realizar.

Arrojados estos datos podemos idear y planificar una serie de objetivos que aporten soluciones a la gente afectada por este tipo de discapacidad.

Los objetivos elegidos son:

-Realizar una aplicación que proporcione un tipo de comunicación inaccesible de otra manera para un invidente. En este caso se ha optado por comunicación corporal facial. La aplicación será capaz de sustraer información facial del interlocutor situado enfrente del dispositivo y entregársela al invidente en formato voz.

-Dada la información presentada anteriormente relativa a los rangos de edad, la aplicación debe ser sencilla y de fácil manejo, ya que las personas más afectadas son de avanzada edad. Existen herramientas de accesibilidad de serie en los sistemas operativos móviles (se explicarán en 3.2) que permiten el uso ordinario del dispositivo, sin embargo estas son tediosas y de difícil manejo. Por tanto se procurará no depender de dichas herramientas.

-El informe de la OMS nos indica en qué idiomas debe estar traducida la aplicación, optando finalmente por italiano, francés, español, inglés, alemán, portugués, ruso y chino.

-Debido a que en las citadas regiones se sitúan muchos países en vías de desarrollo, sería óptimo que la aplicación no requiriese un hardware con características técnicas muy elevadas, tomando así como margen inferior la versión 4.4 de Android.

1.2 Recursos utilizados

-A continuación se definen los recursos tanto de hardware como de software utilizados para la elaboración del proyecto.

- Hardware:
 - Dispositivo móvil ZTE Blade L5
 - Cámara de 8 megapíxeles
 - Procesador de 2 núcleos a 1.3GHz
 - 1GB de memoria RAM
 - Ordenador
 - CPU AMD Phenom II X4 960T 3.00GHz
 - 8GB de memoria RAM
 - Periféricos
 - Cable USB-MicroUSB
- Software:
 - Android Studio (v1.1 y 2.1)
 - Android SDK Tools
 - Android SDK Platform 5.0.1 y 4.4.2
 - ARM EABI v7a System Image (acelerador para CPU AMD)
 - API de Afectiva Toolkit

1.3 Fases de desarrollo

-En este y cualquier proyecto es de vital importancia un planteamiento de las fases de desarrollo que se van a ejecutar. Un proyecto mal planteado en torno a su gestión, desarrollo, delimitación de objetivos y plazos de cumplimiento está destinado al fracaso.

Para ello hay que dividir los problemas distintos problemas de menor alcance e ir resolviéndolos paulatinamente.

En **ANEXO A: Fases de desarrollo** se muestra el WBS (Work Breakdown Structure) empleado para desglosar el proyecto TFG en distintas etapas y sub-etapas.

Dicho WBS se compone de 4 etapas principales destinada en cada caso a tareas de aspectos distintos.

- **Etapas 1.1 Documentación**

- **1.1.1 Investigación ‘Emotion Recognition’:** Documentación e investigación relativa a mecanismos de reconocimiento de emociones y análisis de sentimientos (*‘emotion recognition’* y *‘sentiment analysis’*). Estudio de las clases, elección y búsqueda de herramientas para su desempeño.
- **1.1.2 Estudio de discapacidad:** Posterior a la elección, estudio de fines posibles, entre los cuales se selecciona un fin de accesibilidad para discapacitados. Estudio de gadgets y herramientas existentes de accesibilidad. Test de dichas herramientas.
- **1.1.3 Estudio de plataforma:** Investigación del panorama actual del sistema operativo Android. Estudio de software necesario, actualizaciones y requisitos hardware necesarios para su funcionamiento.
- **1.1.4 Estudio de recursos:** Listado de recursos necesarios y disponibles. Comprobación de su correcto funcionamiento y preparación para su uso en modo desarrollador.
- **1.1.5 Estudio del estado del arte del campo:** Investigación del mercado de aplicaciones destinadas a personas invidentes. Estudio de herramientas de accesibilidad tipo Google Talk.

- **Etapas 1.2 Planteamiento**

- **1.2.1 Elaboración plan de proyecto:** Elaboración de itinerario de desarrollo.
- **1.2.2 Delimitar objetivos:** Definición y acotación de objetivos en base a los recursos y capacidades disponibles.
- **1.2.3 Definir trabajo futuro:** Desarrollo de un prototipo de itinerario excedente que amplíe los objetivos definidos

- **Etapla 1.3 Elaboración**

- **1.3.1 Desarrollo de aplicación:** Elaboración del código necesario para la funcionalidad deseada. Integración de bibliotecas y puesta en funcionamiento. Integración de API Afectiva y Android SDK Tools.
- **1.3.2 Comprobaciones de compatibilidad:** Comprobación de funcionamiento en distintas versiones de sistema operativo simulado en IDE.
- **1.3.3 Plan de pruebas:** Elaboración de un sistema no intensivo de plan de pruebas y ejecución del mismo.
- **1.3.4 Corrección de errores:** Corregir el código fuente en aquellos aspectos erróneos descubiertos en el plan de pruebas.
- **1.3.5 Evaluación:** Evaluación del funcionamiento de la aplicación.

- **Etapla 1.4 Finalización**

- **1.4.1 Estructurar memoria:** Elaboración de tabla de contenidos.
- **1.4.2 Redacción de memoria**
- **1.4.3 Entrega y presentación**

1.4 Estructura de la memoria

-A continuación se realizará una breve descripción del contenido de cada capítulo de la presente memoria.

Capítulo 1 INTRODUCCIÓN: Contiene información sobre la motivación del presente proyecto, así como los objetivos, recursos necesarios y fases para su desarrollo.

Capítulo 2 ESTADO DEL ARTE: Presenta el estado del arte del hardware, software, y profundiza sobre el campo que nos atañe, análisis de sentimientos y reconocimiento de emociones. Describe las capacidades de la herramienta ‘Afectiva’. Muestra también el contexto sobre el que se construye la aplicación, así como el panorama de aplicaciones destinadas para personas con problemas de visión. Incluye finalmente el marco regulatorio pertinente.

Capítulo 3 CICLO DE VIDA DEL PROYECTO: Define las características deseadas y el pronóstico del ciclo de vida de la aplicación.

Capítulo 4 DESCRIPCIÓN DEL SISTEMA: Presenta los requisitos técnicos de software y hardware, y también aquellos aspectos ambientales necesarios para un correcto funcionamiento de la aplicación.

Capítulo 5 FUNCIONAMIENTO DEL SISTEMA: Describe las etapas funcionales de la aplicación. Esto incluye todas las fases a las que el usuario final tendrá que enfrentarse una vez iniciada la aplicación. Finalmente se incluyen comentarios de desarrollador sobre la implementación que enunciará las dificultades acontecidas en la elaboración.

Capítulo 6 EVALUACIÓN DEL SISTEMA: Definición de un sistema de evaluación, así como sus resultados y conclusión sobre los mismos.

Capítulo 7 CONCLUSIÓN: Incluye una conclusión final sobre el proyecto. Análisis objetivo de objetivos logrados y no logrados y análisis subjetivo sobre la satisfacción personal del proyecto.

Capítulo 8 TRABAJOS FUTUROS: Discierne un camino a seguir posterior a la elaboración del proyecto, indicando nuevos caminos o mejoras posibles del mismo.

Capítulo 9 GESTIÓN DEL PROYECTO: Define y presenta la gestión utilizada para la elaboración del proyecto y costes.

Capítulo 10 BIBLIOGRAFÍA

ANEXO A: Fases de desarrollo.

ANEXO B: English Content.

Capítulo 2

ESTADO DEL ARTE

2.1 Sistema operativo: Android

-Android es el sistema operativo elegido para la elaboración de la aplicación y a continuación se presentarán las razones para ello, y más concretamente las versiones utilizadas.

Actualmente el parque de teléfonos móviles a nivel internacional utiliza en su mayoría el sistema operativo Android como muestra la *figura 2.1*:

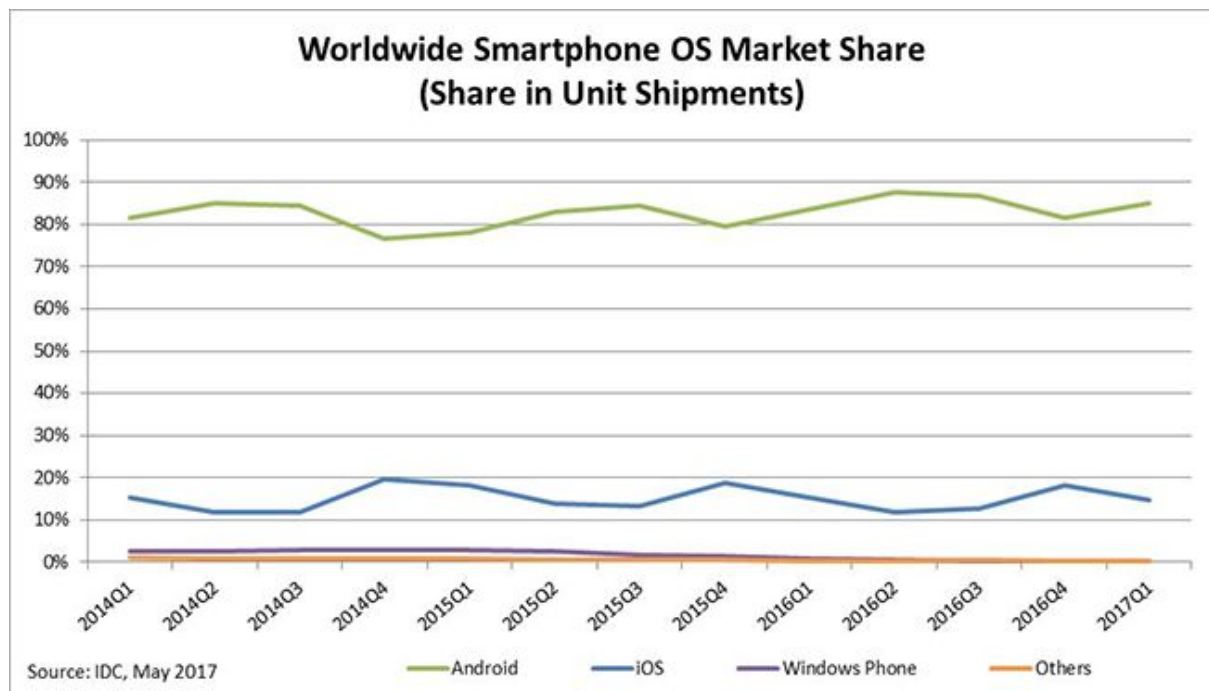


Figura 2.1 Cuota de mercado de sistemas operativos móviles en smartphone [5].

Además, en el último año el porcentaje de dispositivos con sistema operativo Android ha tenido fluctuaciones en torno al 85% de cuota de mercado, siendo esta la actual. Si sumamos esto al historial de cuota en ejercicios anteriores podemos asegurar una estabilidad contundente de su cuota de mercado.

Estos datos equivalen a asegurar que enfocar la aplicación a este sistema operativo nos permitirá tener el mayor alcance posible.

Sin embargo, debido a las características requeridas por la aplicación (explicadas en el capítulo 4), hay que acotar este espectro a las versiones de Android que puedan cumplir con los requisitos. En la *tabla 2.1* se muestra el reparto de versiones de Android. Como se puede observar a partir de la versión

4.4 se impactaría sobre el 91% del parque de móviles con sistema operativo Android. En el presente proyecto se utiliza como target la mencionada versión, sin embargo no ha sido posible realizar los tests oportunos para versiones superiores a 6.0. Además aplicar a este rango de versiones nos proporcionará las características necesarias para un correcto funcionamiento de la aplicación.

Version	Codename	API	Distribution
2.3.3 - 2.3.7	Gingerbread	10	0.7%
4.0.3 - 4.0.4	Ice Cream Sandwich	15	0.7%
4.1.x	Jelly Bean	16	2.7%
4.2.x		17	3.8%
4.3		18	1.1%
4.4	KitKat	19	16.0%
5.0	Lollipop	21	7.4%
5.1		22	21.8%
6.0	Marshmallow	23	32.3%
7.0	Nougat	24	12.3%
7.1		25	1.2%

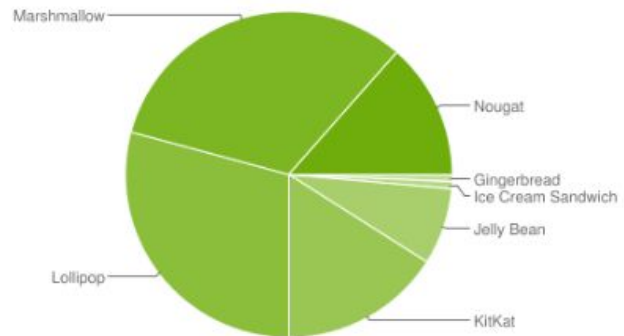


Tabla 2.1 Reparto internacional de versiones Android en 2017 [6].

Pasamos a continuación a destacar brevemente la arquitectura de la plataforma, lo que nos permitirá analizar en mayor detalle el funcionamiento de la misma.

Partiremos desde las capas más inferiores, en las que se procesan los elementos de bajo nivel, y avanzaremos a las capas superiores. La distribución de capas podemos verla en la *tabla 2.2*.

El núcleo (kernel) en el que se basa la plataforma es el de Linux. Contiene los drivers necesarios para el uso de los componentes hardware del dispositivo, como por ejemplo la cámara, teclado, puertos, pantalla, etcétera. Este se debe modificar agregando librerías a medida que se incorporen nuevos elementos hardware al dispositivo.

La capa de abstracción de hardware (HAL) otorga una intermediación estándar entre los componentes hardware y el framework de la API de Java. Se compone de módulos que son utilizados en cada llamada de una API sobre un componente hardware, proporcionando una interfaz que permitirá su uso.

La siguiente capa es el Android Runtime, ésta proporcionará una traducción del bytecode de Java a las instrucciones nativas del dispositivo. Ésta capa ha sido modificada en las versiones recientes del sistema operativo, habiéndose intercambiado la máquina virtual de procesos 'Dalvik' por el Android Runtime (ART).

También se incluyen bibliotecas nativas de C y C++, estas son requeridas por servicios del sistema, como por ejemplo el ART. Otro ejemplo es el OpenGL ES que nos permitirá operar con gráficos 2D y 3D y se utilizará a través de la API Java OpenGL.

En el nivel superior nos encontramos con el Framework de la API de Java. Este framework o entorno de trabajo ofrecerá a los desarrolladores una simplificación para el uso de los diferentes componentes y servicios del sistema. Dentro de los módulos y servicios se incluyen **administradores de recursos**: como gráficos, strings localizados, recursos de diseños; **administrador de notificaciones**, **administrador de actividad**, que administra el ciclo de vida de una aplicación y **proveedores de contenido**, que permite la comunicación de datos entre aplicaciones.

Para finalizar el sistema ofrece unas aplicaciones básicas para el uso del dispositivo, como un navegador de internet, una aplicación para el uso de la cámara, el teclado, etcétera. Sin embargo estas aplicaciones no poseen un estatus especial sobre cualquier otra aplicación descargada de la ‘store’ de Android.

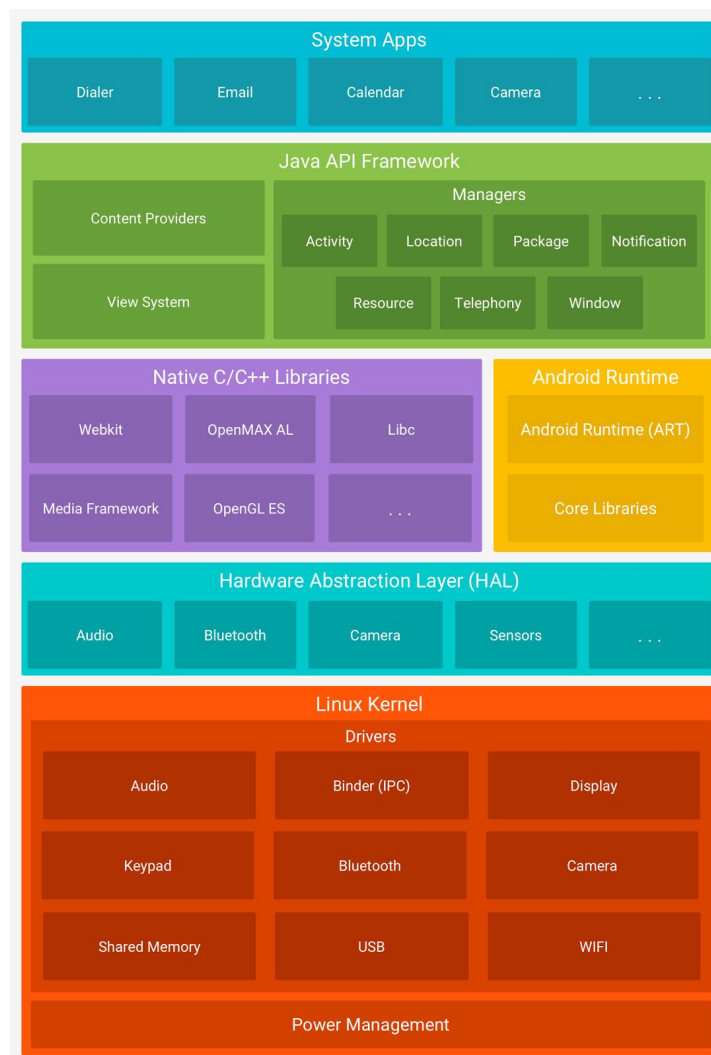


Tabla 2.2 Arquitectura de software de Android [6]

2.2 Texto a Voz

-Las herramientas disponibles de TTS (Text to speech) disponibles actualmente para la plataforma son mucho más accesibles y completas que en versiones anteriores. Android nos ofrece bibliotecas dentro de la API de Android a partir de la versión 4, sin embargo se han ido mejorando y agregando funcionalidades hasta las versiones actuales.

Es posible sintetizar voz a partir de texto para una reproducción inmediata o creación de archivos de sonido. Lo único necesario para su implementación son las bibliotecas de idiomas para cada lenguaje determinado. En los dispositivos comercializados en territorio europeo los idiomas disponibles en la aplicación de ajustes son los siguientes:

- | | |
|------------------|------------|
| ● Finlandés | ● Danés |
| ● Eslovaco | ● Holandés |
| ● Chino mandarín | ● Polaco |
| ● Rumano | ● Noruego |
| ● Croata | ● Francés |
| ● Serbio | ● Griego |
| ● Inglés | ● Turco |
| ● Portugués | ● Español |
| ● Húngaro | ● Italiano |
| ● Alemán | ● Ruso |
| ● Sueco | ● Checo |
| ● Búlgaro | |

(No se han contemplado lenguajes comunes a distintas geolocalizaciones)

2.3 Análisis de sentimientos y reconocedor de emociones

-Gracias a los avances en la velocidad de computación y cálculo de los procesadores de las últimas décadas ha sido posible poner al alcance del consumidor algoritmos que nos permiten calcular las sensaciones subjetivas de una persona con respecto a una entrada de información. Es posible deducir el estado emocional que induce en una persona un *input* presentado en distinto formato (video, fotografía o texto), o también descifrar computacionalmente aspectos de carácter subjetivo sobre un texto, como por ejemplo su positividad o negatividad, si contiene ironía, o la índole del mismo, deduciendo si se trata de un texto expositivo, una crítica, etcétera; así como el tema del que trata.

A continuación se presentarán ejemplos de analizadores de sentimientos y reconocedor de emociones que se aplican sobre distintos formatos de entrada. Se explicará brevemente lo que ofrecen y los posibles usos inmediatos en los que se aplican.

❖ **Sobre texto:**

En este grupo se incluyen herramientas como IBM Watson, Text Analytics de Microsoft, Text2data, Natural Language Tool Kit (NLTK), Theysay, entre muchos otros.

Metodología:

-Existen varios tipos de procesos para llevar a cabo un análisis de sentimientos basado en una entrada en formato texto [7]. A continuación se describirán brevemente algunos de ellos:

- *Método de localización de palabras clave:*
Es un método fácil de implementar e intuitivo ya que únicamente consiste en localizar palabras que estén preestablecidas dentro de un rango de sentimientos (tristeza, alegría, enfado, miedo, etcétera). El proceso analiza el texto buscando dichas palabras clave y comprueba el énfasis aplicado, posteriormente da como resultado la emoción obtenida con mayor resultado.
- *Método de afinidad léxica:*
Este método opera al nivel léxico de la palabra, es decir, puntúa cada palabra en base a la emoción relacionada con su significado. El análisis del texto consiste en analizar cada palabra y clasificarlas según la emoción que evoca independientemente. El problema de este método reside en que no se tiene en cuenta el contexto en el que se sitúa cada palabra, por lo tanto puede dar lugar a incorrelaciones.
- *Métodos de aprendizaje:*
Este tipo de métodos intentan reconocer las emociones de un texto aplicando un clasificador previamente entrenado, el cual ha identificado textos, frases, palabras y partículas dentro de una categoría de emoción para posteriormente aplicarlo a un conjunto de textos de test.

Casos de uso:

- Monitorizado de la audiencia y escucha de la opinión de la sociedad. Es posible monitorizar en tiempo real las opiniones sobre un producto después de un evento mediante sus opiniones en las redes sociales.
- Marketing personalizado en base a las emociones del cliente. Otorga la posibilidad de modificar los mensajes publicitarios en función del estado de ánimo del cliente proporcionando así más empatía en el anuncio, lo que ayudará a su asimilación e influencia.

- Personalización en mensajes automatizados. Posibilidad de analizar los mensajes en una interacción con el cliente y construir una respuesta acorde a la situación.
- Análisis automatizado de textos. Otorga la capacidad de clasificar y sintetizar textos, por ejemplo de índole periodística, según el tema del que tratan y la positividad o negatividad plasmada en las opiniones del texto.

❖ **Sobre voz:**

Dentro de esta categoría se incluyen aplicaciones como las de ‘Empath’ de Smartmedical Corp., ‘Vokaturi’ o ‘Good Vibrations’

Metodología:

-Lo métodos utilizados para este grupo de reconocimiento de emociones es independiente al contenido de lo que se dice en los mensajes orales, es decir, el análisis recaerá en el tono, frecuencias, velocidad, etcétera, y no en el significado o idioma que contenido en el discurso. Los factores sonoros a utilizar generalmente son el tono de la voz, la velocidad del habla, el volumen y los distintos eventos que ocurren en un discurso, como por ejemplo sílabas, fonemas explosivos, nasales, o también los silencios [8][9]. El proceso para analizar todos estos factores y detectar la emoción en un discurso se basan en redes de aprendizaje, clasificadores estadísticos que son entrenados y posteriormente probados.

Casos de uso:

- Monitorización y control del feedback obtenido durante las llamadas telefónicas en servicios de atención al cliente. Es posible cuantificar y gestionar información relativa a la satisfacción del soporte de ayuda o promocional proporcionado a un cliente mediante servicio de voz.
- Personalización de mensajes en la interacción entre persona-máquina, gracias a la información obtenida de los rasgos emocionales incluidos en el discurso.

❖ **Sobre imagen facial:**

En este grupo se centra la API que se usará en este proyecto, Affective, pero también existen otras aplicaciones como Emotion API de Microsoft, EmoVu, Kairos o FaceReader de Noldus, entre otros.

Metodología:

-Las expresiones faciales otorgan mucha información acerca de las emociones que una persona siente. Elementos como las cejas, los párpados, la boca y las mejillas nos brindan la oportunidad de conocer el estado de ánimo de alguien sin necesitar mediar palabra. Es por ello

que los métodos computacionales para obtener esa información se basan en la configuración de la distribución espacial de determinados puntos faciales de interés.

En una primera etapa se localizan dichos puntos utilizando técnicas de tratamiento de imagen y modelos geométricos. Posteriormente se realiza un seguimiento de cómo esos puntos se desplazan en la imagen, como se muestra en la *figura 2.2*, para así tomar esos datos y aplicarlos a un sistema de clasificadores que nos dará la probabilidad de a qué emoción se asemeja más. Esto se ayuda de movimientos faciales simples, como por ejemplo un levantamiento de las cejas, lo cual tendrá en este caso más probabilidad de pertenecer a una emoción de sorpresa o miedo que de otras emociones como tristeza. Juntando la información de todos los movimientos musculares se puede esbozar una pronosticación sobre la emoción.

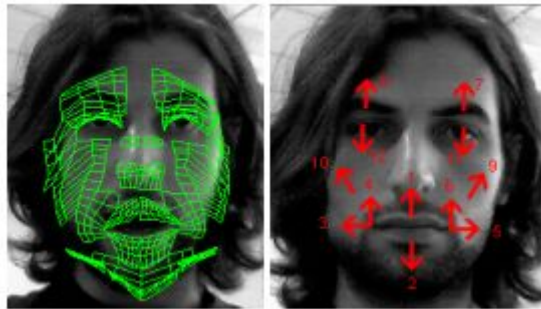


Figura 2.2 Previsualización de algoritmos de rastreo de expresiones faciales [10].

Casos de uso:

- Comprobación de la efectividad de contenido promocional sobre el posible cliente. Es posible diagnosticar la efectividad de un anuncio u oferta de productos sobre el consumidor analizando su respuesta facial en el momento del visionado. Esto permitirá construir una estrategia de marketing adecuada efectiva.
- Medición del grado de satisfacción ante eventos en contenido audiovisual. En base a la respuesta obtenida por los espectadores de un contenido audiovisual es posible obtener información sobre el grado de satisfacción de dicho evento.
- Interactividad entre jugador y juego en función de los gestos faciales. Esto proporciona un nivel más de interactividad sobre el que construir experiencias videojugabilísticas.
- Facilidad en interacciones máquina-humano. Cualquier dispositivo que contenga una entrada de información visual, y concretamente facial, del usuario, le podrá permitir obtener información de sus emociones para poder así tomar determinadas acciones en base a ello. Esto se puede aplicar, por ejemplo a la robótica.

2.4 Afectiva

-Afectiva será la herramienta utilizada para llevar a cabo nuestra aplicación y a continuación se explicarán sus características en detalle.

Como se ha explicado anteriormente, Afectiva sustrae información realizando un reconocimiento emocional a partir de la imagen facial de una persona, procesándola mediante técnicas de tratamiento de imagen y aprendizaje máquina.

Las medidas que puede sustraer la herramienta se pueden agrupar en:

- ❖ **Apariencia**
- ❖ **Emociones**
- ❖ **Expresiones faciales**
- ❖ **Emojis**

- ❖ **Apariencia:**

-En este conjunto de resultados se proveerán aquellos elementos relacionados con el aspecto físico, como son la edad, la etnia, el género y si lleva o no gafas.

- **Edad:**

-Sobre este aspecto se puede llegar a estimar la edad del individuo clasificándola en rangos de edades comprendidas en:

- Menor de 18 años.
- Entre 18 y 24 años.
- Entre 25 y 34 años.
- Entre 35 y 44 años.
- Entre 45 y 54 años.
- Entre 55 y 64 años.
- Mayor de 65 años.

- **Etnia:**

-El clasificador de etnia clasifica a la persona en uno de los siguientes grupos mostrados en la *tabla 2.3*:

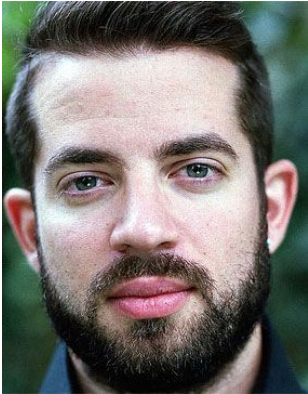



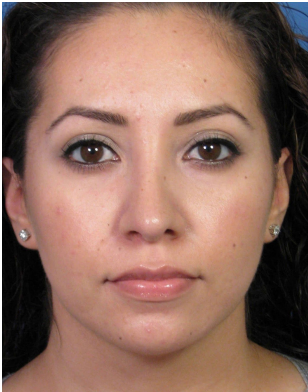
-Caucásico	-Negro africano	-Sur de Asia
		
-Este de Asia	-Hispano	
		

Tabla 2.3 Etnias del identificador de etnia de Affective.

- **Género:**

-El clasificador de género intenta asemejar la percepción humana relativa al género de una persona. Definiendo, obviamente, a una persona como *masculino*, *femenino* o *desconocido* en caso de no ser capaz de dar un valor relativamente seguro.

- **Gafas:**

-Nos ofrecerá información sobre si la persona posee gafas de ver, gafas de sol o si no lleva gafas.

- ◆ **Emociones:**

-Los resultados de esta categoría nos ofrecerá las similitudes a las siete emociones básicas a partir de las expresiones faciales universales, que se muestran en la *tabla 2.4* [11]:

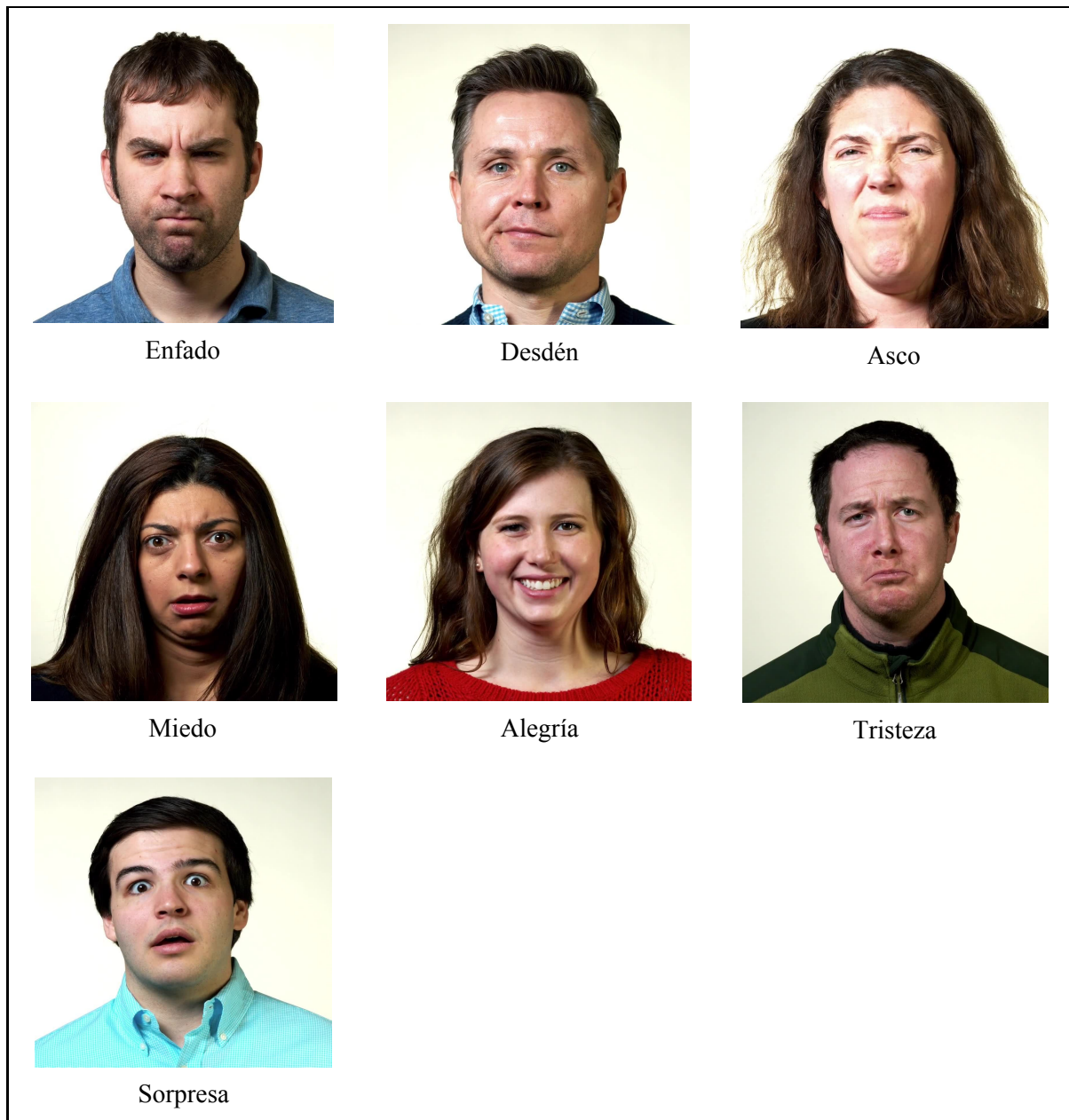


Tabla 2.4 Emociones del identificador de emociones de *Afectiva* [12].

❖ Expresiones faciales

-La herramienta también nos permite identificar distintas expresiones faciales hasta un total de 21 diferentes. Estas expresiones son más básicas que las que podrían ocurrir al realizar una emoción básica. Comprenden movimientos de los músculos faciales de frente, cejas, nariz, mejillas y boca. Esta es la lista de expresiones disponibles:

- Atención
- Levantar cejas
- Ceño fruncido
- Levantar las mejillas

- Levantando la barbilla
- Hoyuelos
- Ojos cerrados
- Ojos bien abiertos
- Levantamiento de la parte interna de las cejas
- Mandíbula abierta
- Ojos achinados
- Comisuras de los labios caídas

- Labios oprimidos
- Labios en forma de beso
- Labios estirados hacia atrás.
- Labios succionados
- Boca abierta
- Nariz arrugada
- Sonrisa
- Sonrisa afectada (Smirk)
- Labio superior levantado

◆ Emojis

-Podemos también reconocer e identificar expresiones faciales a emojis por todos conocidos, para simplificar y comunicar de manera más simple y directa la información que se obtiene con el reconocedor emocional. En la *tabla 2.5* se muestran las distintas expresiones:














		
Riendo	Sonriendo	Relajado
		
Guiñando un ojo	Mandando un beso	Sacando la lengua
		
Guiñando un ojo y sacando la lengua	Gritando	Sonrojado
		
Sonrisa afectada (smirk)	Decepcionado	Enfadado
		
	Neutral	

Tabla 2.5 Emojis del identificador de emojis de Afectiva.

2.5 Panorama de aplicaciones para invidentes

-A continuación se mencionan algunas aplicaciones y tecnologías destinadas para personas invidentes [15]. Esto nos permitirá saber qué posibilidades están al alcance de una persona con ceguera para mejorar en algún aspecto su calidad de vida, y conocer también si este proyecto podría encajar en lo que solicita el mercado.

- **FingerReader [16]:**

-Este dispositivo asiste a invidentes y personas visualmente discapacitadas en la lectura de texto impreso. El aparato se coloca en el dedo índice, como vemos en la *figura 2.3*, y posee una cámara que recibirá la imagen del texto impreso a medida que el usuario vaya siguiendo las líneas de texto. Mediante algoritmos de tratamiento de imagen el aparato proporcionará una salida de audio transformando el texto impreso en voz, la cual se enviará al usuario.

Debido a los posibles errores en el seguimiento de la línea, el portador recibirá retroalimentación que le permitirá conocer si se ha finalizado o comenzado un renglón, o si su dedo ha variado la trayectoria de la línea en el movimiento.

FingerReader aún se encuentra en fase de investigación y es un prototipo actualmente, pero se espera desarrollar el producto final próximamente. Está desarrollado por el MIT Media Lab.

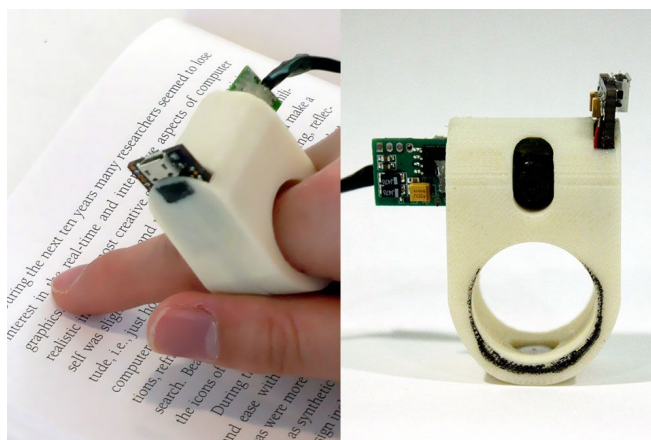


Figura 2.3 Dispositivo para la asistencia de lectura FingerReader.

- **KNFB Reader [17]:**

-Para un uso similar al de FingerReader, que un invidente tenga la capacidad de leer, pero con una metodología distinta, disponemos en las tiendas de apps la aplicación KNFB Reader. Al iniciar la aplicación se ejecutará la cámara, con la cual se deberá realizar una fotografía a aquello que se desee leer, como se ve en la imagen de la *figura 2.4*. Posteriormente la aplicación operará con esa imagen para obtener el texto que aparece en ella y transformarlo en mensajes de voz que se emitirán por la salida de audio.

La ventaja que nos ofrece esta aplicación radica en la posibilidad de leer textos en cualquier formato, ya que es posible fotografiar textos situados en carteles, monumentos, pantallas, señales y cualquier otro texto impreso tanto en papel como en objetos.

Cabe destacar también su practicidad, ya que permite hacer del móvil un dispositivo capaz de otorgar la posibilidad de leer a un invidente simplemente con una aplicación.

La aplicación está disponible en la AppStore, GooglePlay y la tienda de Windows 10, con compatibilidad para iOS, Android y Windows 10. Está desarrollada por Sensotec y KNFB Reader LLC.



Figura 2.4 Captura del video promocional de KNFB Reader [18].

- **Color ID [19]:**

-Otra aplicación que ofrece a invidentes capacidades básicas para el resto de personas es Color ID. Gracias a esta aplicación es posible que una persona con ceguera sea capaz de saber qué colores le rodean. La aplicación iniciará la cámara y reproducirá por la salida de audio los colores que se enfocan. Una aplicación muy sencilla pero que facilita la vida a los invidentes ayudándoles en tareas cotidianas como por ejemplo combinar los colores de la vestimenta o diferenciar entre objetos con la misma forma y distinto color.

- **TapTapSee [20]:**

-Esta aplicación móvil es un identificador de objetos. Está diseñada para que personas invidentes reciban ayuda identificando objetos con ayuda del teléfono móvil. Al iniciar la aplicación entra en funcionamiento la cámara, con la cual el invidente deberá enfocar aquello que quiera reconocer y pulsar dos veces la pantalla. Posteriormente se emitirá por el altavoz (utilizando Google TalkBack en la versión para Android) una descripción de la imagen capturada.

Es capaz, por ejemplo, de identificar billetes o el género de una persona. Y es que sorprende la eficacia del resultado obtenido ya que describe imágenes relativamente complejas y compuestas por varios elementos con solvencia, incluso aportando cierto detalle como el color de los objetos o el escenario. Esto se explica porque la aplicación se ayuda de una combinación de algoritmos de búsqueda y decisión, y un sistema de colaboración abierta.

La imagen capturada es procesada durante unos segundos mientras se realiza el análisis. Éste puede producirse de dos maneras diferentes:

-Mediante un tratamiento de la imagen para un posterior algoritmo de búsqueda en las bases de datos con las que opera la aplicación.

-Mediante la descripción proporcionada por una persona física que observará la imagen y la describirá sintetizadamente, para devolverla como resultado.

Se desconoce cuándo se opta por un sistema u otro al analizar la imagen, por lo que se recomienda no fotografiar información sensible y/o personal.

En definitiva es una aplicación muy útil para las personas invidentes, pues simplifica mucho las tareas cotidiana y les otorga más independencia.

- **Be my eyes [20]:**

-Esta aplicación tiene un funcionamiento muy sencillo. Proporciona asistencia visual a un invidente o discapacitado visual simplemente pulsando la pantalla. La aplicación posteriormente conectará con un voluntario que recibirá una notificación y se establecerá una conexión de video y voz entre los dos. El voluntario (sin discapacidad visual) explicará en la llamada los objetos a los que la cámara está apuntando y proporcionará ayuda al discapacitado en caso necesario. Además, la aplicación se encarga de que el usuario conecte con un voluntario que hable su mismo idioma y que habite en una franja horaria oportuna para no molestar.

Desde el punto de vista del voluntario, le llegará al móvil una notificación anunciando que un usuario de la aplicación requiere asistencia. Si aceptan la llamada se iniciará la comunicación de video y voz. Si el voluntario rechaza o ignora la llamada, esta se trasladará a otro voluntario, así sucesivamente hasta que se responda.

No ocurren largas demoras y es posible conseguir asistencia a cualquier hora del día debido a la cantidad de voluntarios activos [21].

En la praxis el discapacitado adquiere independencia al no requerir de asistencia presencial. Los usos pueden ser muy variados, desde asistencia para identificar artículos en el supermercado a saber si se está haciendo bien la colada, juntando la ropa de color o tipos de tejido.

2.6 Marco Regulatorio

-La aplicación no guardará en el terminal datos de contenido sensible que puedan afectar a la privacidad del usuario, acorde con la Ley Orgánica 15/1999 de Protección de Datos de Carácter Personal [13].

El acuerdo con Afectiva no permite revender, redistribuir sus aplicativos o los reportes de datos creados por sus productos para ningún propósito, salvo únicamente aquellas ocasiones en las que se permita un acuerdo distinto con Afectiva. Se acepta también no generar ningún ingreso obtenido a partir de un producto estándar o de desarrollo que pueda considerarse razonablemente similar a uno de sus aplicativos. A su vez el acuerdo no permite dañar, deshabilitar o congestionar los servidores o redes conectados a los sistemas de Afectiva.

La licencia de desarrollo no permite desarrollar aplicaciones o productos diseñados para vulnerar la seguridad o intimidad de personas sin su consentimiento. No se permite tampoco lanzar un producto cuya licencia sea ‘open source’, ‘copyleft’ o ‘community source’.

Además no se permitirá con esta licencia desarrollar productos diseñados para un entorno político, de investigación de marketing o investigación académica.

Por su parte Afectiva asegura cumplir todas las leyes y regulaciones aplicables relativas a la privacidad y al posible uso de imágenes o videos de usuarios o cualquier información identificable de un usuario.

Con respecto a la patente de la aplicación resultante, esta práctica se excluye explícitamente en el artículo 52 párrafo 2 del Convenio de Patente Europea sobre programas de ordenador [14]. Solamente serán patentables aquellas invenciones que solucionen técnicamente un problema técnico, como por ejemplo un software o un protocolo que permita un mayor rendimiento en comunicaciones telemáticas. Lo cual requiere buscar otras estrategias para poder reservar los derechos de propiedad en la medida de lo posible. Estas alternativas irán enfocadas a registrar las marcas, las interfaces, y la autoría del código utilizado.

Capítulo 3

CICLO DE VIDA DEL PROYECTO

3.1 Características de la aplicación

-Una vez analizado el panorama tanto de la tecnología de analizadores de sentimientos como de las aplicaciones y tecnologías destinadas a personas con discapacidad visual podemos sacar una serie de conclusiones que nos permitirán sustraer ideas de diseño para la aplicación.

La clase de analizador de sentimientos o reconocedor de emociones escogido será el de expresiones faciales ya que la información que se comunica en este campo es de naturaleza corporal y visual. Los analizadores relativos a la voz carecen de sentido para ser usado por personas físicas cuya capacidad les permita obtener la información que proporciona esta tecnología. Los analizadores que trabajan con textos sí serían favorables para usuarios con capacidad visual reducida, sin embargo existen, como hemos visto en 3.1 muchas aplicaciones relacionadas con la lectura de textos, y la obtención de información de carácter emocional o sentimental de los mismos no tiene relación con la incapacidad visual, por tanto no sería oportuno a simple vista utilizar dichos sistemas.

En conclusión la rama utilizada de reconocedor de emociones será la que trabaja con expresiones faciales ya que proporcionará información de carácter emocional sobre información representada visualmente y que requiere procedimientos de lectura que no han sido utilizados en aplicaciones móviles hasta ahora. Además puede proporcionar información de carácter descriptivo del físico de la persona a través del aspecto facial. Esto otorgará un campo de la información que todas las personas sin discapacidad visual posee y que se ve privada en aquellas con invidencia.

En relación a las aplicaciones móviles en el mercado destinadas a personas invidentes o con discapacidades visuales podemos sacar las siguientes ideas sobre aspectos comunes entre ellas:

- **Simpleza:** Debido a lo lento que puede resultar navegar por los botones, pantallas y textos utilizando la herramienta de accesibilidad Google Talka, la cantidad de botones y pantallas es escasa. Esto contribuye a que la aplicación sea de uso fácil y rápido.
- **Wearable (adaptable a wearable):** La característica principal de la tecnología wearable es precisamente su alta frecuencia de uso. Portar un objeto con esta tecnología va a significar que ese objeto va a ser utilizado intermitente o constantemente. Ello implica que el uso de la aplicación deberá ser adaptable a un wearable o utilizado con una autonomía suficiente para que no requiera una interacción constante con el dispositivo para su uso. Es decir, una independencia en su funcionamiento notable.

- **No interrumpe la acción del usuario. Complemento:** Es importante también que la aplicación no sea invasiva y no interrumpa el ritmo de las acciones del momento. Como cualquier otra aplicación o dispositivo de accesibilidad, se debe pretender parecerse lo máximo posible a la nula notabilidad, es decir intentar hacer como si no existiera. En las aplicaciones enunciadas se realiza interacciones digitales utilizando Google Talk para navegar por las opciones de la aplicación, sin embargo estas procuran ser lo más breves posibles para no pausar prolongadamente la acción llevada a cabo en ese momento.
- **Independencia:** En relación a las finalidades de las aplicaciones, todas aportan para otorgar independencia al usuario. Se intenta depender de un dispositivo móvil y no de otra persona, lo cual aumenta el grado de independencia. Estas tareas van a ser simples (como simplemente poder leer o saber colores de objetos).
- **Google Talk:** Las aplicaciones se apoyarán si es necesario en la herramienta de accesibilidad Google Talk para navegar por menús. Como se ha mencionado anteriormente, esta herramienta es tediosa y lenta en su uso ya que requiere navegar por todos los botones que aparecen en la pantalla sucesivamente para llegar a la opción deseada (en 4.1.1.2 se explicará su funcionamiento)
- **Interfaz gráfica básica:** En su diseño gráfico, las aplicaciones móviles para invidentes van a poseer escasos elementos decorativos debido a la incapacidad de su disfrute visual por parte del usuario.

3.2 Ciclo de vida del proyecto

-Teniendo ya las ideas principales sobre la aplicación se debe proceder a llevar a cabo una metodología que permita utilizar de manera efectiva los recursos materiales y temporales.

Debido a la independencia existentes en las partes del trabajo final se ha optado por modular los diferentes hitos del proyecto separándolos en función de la funcionalidad del proyecto.

En primer lugar se procederá a tareas de documentación para obtener información sobre los temas a tratar. Se hace especial atención a los siguientes campos:

- ‘Sentiment analysis’ y ‘Emotion recognition’
- Traductores Text to Speech y otros sistemas de conversiones para accesibilidad. Investigación a su vez de diferentes motores de conversión.
- Documentación relativa a la utilización de nuevas versiones de Android, uso y diferencias con versiones anteriores.

Posterior a la documentación se llevan a cabo tareas de delimitación de objetivos. En todo proyecto este paso es fundamental. Conociendo la meta del proyecto se evita caer en excesivas ramificaciones en el planteamiento que pueden surgir en su desarrollo. Realizar un plan de hitos y objetivos es crucial para cumplir con los tiempos y concentrar los recursos en aquello realmente necesario. Además, un buen planteamiento de objetivos y procesos nos permitirá realizar un plan de trabajo que beneficiará para adelantar tareas y saber el estado en el que se encuentran los procesos sin terminar.

En el desarrollo, son notables los grupos con funcionalidades diferenciadas que formarán después parte de un todo. Estos son:

- Utilización de Text to Speech
- Utilización de descriptor facial
- Diseño de interfaz
- Integración de todas las partes

Una vez desarrolladas todas las partes se realizará un pequeño control a modo de evaluación del sistema que nos ayudará a calibrar y realizar ajustes finales en diferentes parámetros.

Capítulo 4

DESCRIPCIÓN DEL SISTEMA

4.1 Descripción de la aplicación

-La aplicación será una herramienta destinada para personas con ceguera o problemas severos de visión, cuya misión consistirá en poder transmitir al invidente información de la comunicación corporal, concretamente facial, de su interlocutor.

El análisis de aspecto y la detección de emociones y expresiones faciales se basará en las imágenes obtenidas en tiempo real de la cámara frontal o trasera, dependiendo de la situación, del teléfono.

La transmisión de la información desde la herramienta al usuario se realizará por medio de la salida de audio, utilizando recursos de síntesis de voz. Su procedimiento se basará en la generación de diferentes frases dependiendo del aspecto o expresiones eventuales que el interlocutor haga, para posteriormente ser reproducidas por el sintetizador de voz y que así el teléfono ‘hable’ al usuario.

La información que la aplicación obtendrá del interlocutor será:

- Apariencia:
 - Sexo
 - Rango de edad
 - Etnia
- Emociones:
 - Tristeza
 - Alegría
 - Enfado
 - Sorpresa
 - Desdén
- Eventos:
 - Sacar la lengua
 - Sacar la lengua y guiñar un ojo
 - Guiñar un ojo
 - Abrir la boca
 - Cerrar los ojos
 - Morderse los labios
 - Abrir plenamente los ojos
 - Enfadarse
 - Entristecerse
 - Sorprenderse
 - Mostrar desprecio

- Sonreír
- Sonreír mostrando los dientes

Los mensajes reproducidos por la aplicación no serán intrusivos ni persistentes, se deberá respetar tiempos de espera para no fatigar ni interrumpir la acción del usuario. Se llevará a cabo una función pasiva y complementaria a la acción del invidente.

Además, la aplicación es traducida a 8 idiomas distintos para poder llegar así a un número considerable de usuarios. El lenguaje utilizado es muy sencillo y descriptivo.

Los ámbitos de uso en los que se puede aplicar la aplicación son, por ejemplo:

- Ámbito laboral:

En aquellas tareas donde sea necesario tratar de cara al público o entrevistando personas y donde la conversación sea algo importante en la tarea, como por ejemplo una entrevista, negociación o una clase. O también en los contextos donde la descripción física tome partido en la finalidad laboral, como por ejemplo en encuestas de valoración de un servicio.

- Ámbito de ocio o vida cotidiana:

Puede ser aplicable en la interacción con familiares o amigos, otorgando la posibilidad de saber qué expresiones faciales ocurren, por ejemplo, en los silencios de una conversación. También como complemento de contenidos multimedia donde aparezca la imagen facial del creador de contenidos y ello sea un aspecto importante de su producto, como por ejemplo streamings o videos de personas jugando a videojuegos.

El escenario de aplicación de la herramienta requiere algunos requisitos indispensables referidos a las necesidades técnicas y ambientales. Sin ellos se disminuye la calidad de la experiencia del usuario.

4.2 Especificación de requerimientos

-En este capítulo se enunciarán los requerimientos necesarios para la ejecución de la aplicación y su correcto funcionamiento. Estos requerimientos no solo incluyen aspectos técnicos y materiales pero también las características ambientales referidas a la luz necesarias para conseguir un resultado óptimo.

4.2.1 Requerimientos técnicos

A continuación se enuncian los requisitos técnicos recomendados del dispositivo:

4.2.1.1 Dispositivo

- Procesador Dual-core, 1.5 GHz Cortex-A53.

- 1 GB de memoria RAM.
- Dispositivo capaz de ejecutar API Android 16 o superior.
- Java 1.7 o superior para el dispositivo móvil empleado para el desarrollo.

4.2.1.2 Google TalkBack

Este servicio ofrece la posibilidad de controlar el dispositivo completamente sin necesitar la interfaz visual. Ayuda entonces a los usuarios invidentes o con discapacidad visual añadiendo mensajes de voz o vibraciones para indicar y leer los que ocurre en la pantalla del teléfono.

Este servicio suele estar instalado de fábrica en muchos modelos de dispositivos, sin embargo esto depende de varios factores como por ejemplo el fabricante, la versión de Android y la versión de Talkback. Para iniciar este servicio se debe acceder a la aplicación de Ajustes y posteriormente abrir la opción de Accesibilidad. Ahí se encontrará la aplicación TalkBack para activar o desactivar como ve en la *figura 4.1*:

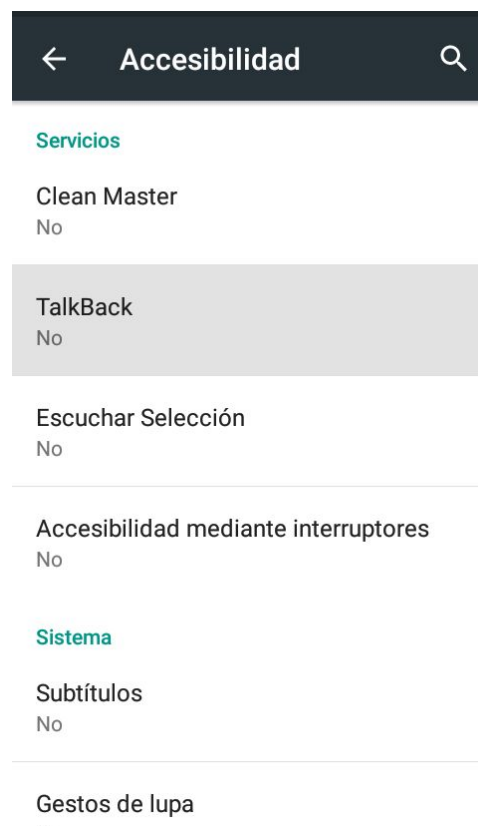


Figura 4.1 Captura de Ajustes/Accesibilidad, donde se podrá activar TalkBack.

Su funcionamiento consiste en realizar gestos con los dedos una vez iniciada la aplicación, para navegar por los textos y botones de las aplicaciones.

Los gestos se dividen en tres grupos: básicos, hacia adelante o atrás y por último gestos en forma de L. En la *tabla 4.1* se clasifican los gestos según las acciones que ejecutan.

Acción	Gesto
Ir al siguiente elemento de la pantalla	Deslizar el dedo hacia la derecha
Ir al elemento anterior de la pantalla	Deslizar el dedo hacia la izquierda
Desplazarse por los ajustes de navegación	Deslizar el dedo hacia arriba o hacia abajo.
Seleccionar el elemento marcado	Tocar dos veces

Tabla 4.1 Tabla de gestos y acciones de Google TalkBack [22].

A medida que se cambie de pantalla o se actualice, por ejemplo añadiendo nuevos elementos, se irán reproduciendo por la salida de audio aquellos eventos para tener siempre información de la interfaz visual.

Los elementos seleccionados y/o pulsados se reproducirán, así como textos, lista de ítems, botones, checkbox, radios... Se aportará información sobre la cantidad de ítems en una lista de elementos y los enunciará sucesivamente (como, por ejemplo, las que aparecen en Configuración). Los textos se leerán automáticamente y será posible leerlos de diferentes modos: carácter a carácter, línea a línea, por párrafos o todo el texto completo. Se cambiará de un modo a otro de lectura realizando un gesto básico después de seleccionar un elemento de tipo texto.

Se pueden adaptar las aplicaciones para mejorar la accesibilidad de TalkBack agregando etiquetas en los elementos de la interfaz. Dichas etiquetas permiten a lectores de pantalla como TalkBack describir con detalle las funciones de los diferentes controles que aparezcan en pantalla. Por ejemplo, aquellas imágenes que a su vez sean botones se deberán etiquetar de forma que el lector de pantalla enuncie cuál es el propósito del botón. También se recomiendan prácticas de diseño para facilitar la navegación en la aplicación a discapacitados visuales, por ejemplo, agrandando los ítems seleccionables, aumentando el ratio de contraste de colores o utilizando diferentes tonos de colores.

4.2.1.3 SO Android

- Para la aplicación se requiere que el dispositivo sea compatible con versiones de Android 4.4 o superior.

4.2.2 Requerimientos ambientales

-Para poder localizar y analizar las expresiones faciales es necesario también cumplir con aspectos ambientales que permitan discernir los puntos de interés de la imagen. Estos aspectos se refieren a la cantidad y tipo de luz que incide sobre la cara de la persona enfocada.

Los resultados obtenidos de los análisis mejorarán cuanto mejor distinga la cámara los detalles de la cara. Por tanto, obviamente, los resultados serán más óptimos en casos de incidencia de luz frontal y a ser posible blanca o natural.

4.2.2.1 Luz natural

-A continuación, en la *figura 4.2* y *figura 4.3*, se muestran, a modo de ejemplo, los resultados obtenidos en condiciones de luz natural frontal y luz natural incidiendo de espaldas respectivamente.

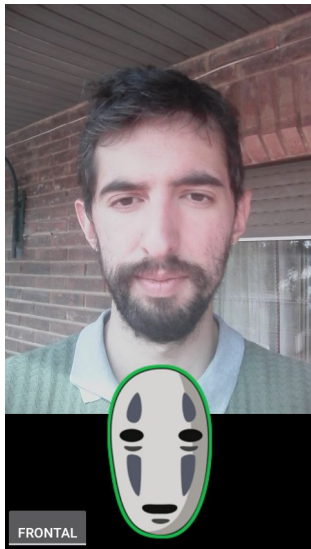


Figura 4.2 Captura de los resultados ante un escenario con luz natural incidiendo frontalmente.

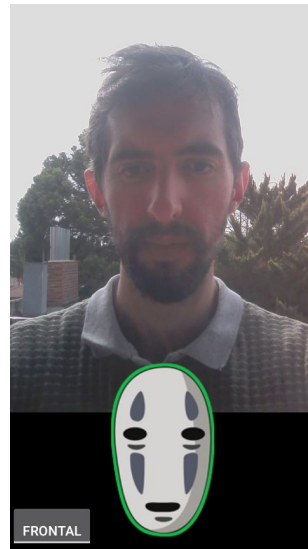


Figura 4.3 Captura de los resultados ante un escenario con luz natural incidiendo de espaldas.

El resultado para la figura 4.2 es el esperado:

“Es un hombre. De entre 25 y 34 años. Y parece de etnia caucásica.”

El resultado para la figura 4.3 es el siguiente:

“Es un hombre. De entre 25 y 34 años. Y parece de etnia hispana.”

Concluimos entonces que la luz puede afectar a los resultados, en este caso tanto de edad como de etnia. Esto se debe a que la tonalidad de la piel, en este caso entre etnia caucásica e hispana, es donde radica la diferencia, y, siendo esta modificada en todos los puntos de la imagen facial provocará una discrepancia.

4.2.2.2 Luz artificial

-En la figura 4.4 y la figura 4.5 se muestran escenarios donde la luz es artificial e incide frontal y lateralmente.

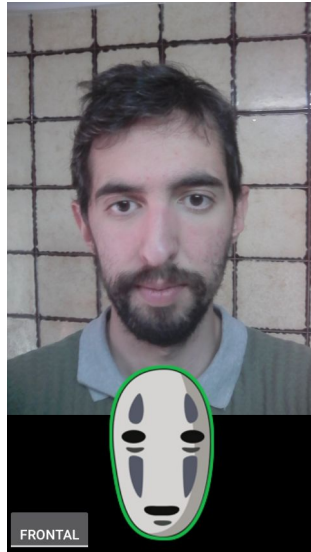


Figura 4.4 Captura de los resultados ante un escenario con luz artificial incidiendo frontalmente.

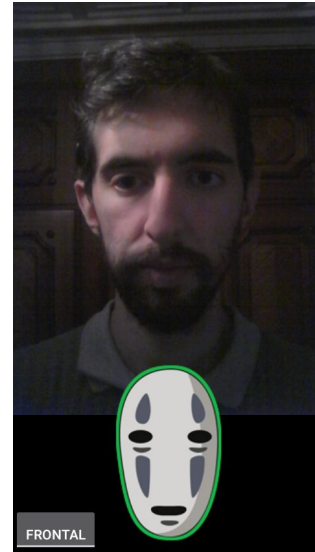


Figura 4.5 Captura de los resultados ante un escenario con luz artificial incidiendo lateralmente.

En el primer caso se obtiene el resultado obtenido:

“Es un hombre. De entre 25 y 35 años. Y parece de etnia caucásica.”

En el segundo caso ocurre discrepancia en la edad:

“Es un hombre. De entre 35 y 44 años. Y parece de etnia caucásica.”

Como era de esperar, la calidad de la luz influirá en la veracidad de los resultados. Es por ello que se recomienda utilizar la herramienta en escenarios donde haya buena iluminación.

Capítulo 5

FUNCIONAMIENTO DEL SISTEMA

5.1 Selección de idioma

-Dado que esta discapacidad no entiende de idiomas, la aplicación está traducida a 8 lenguajes, incluyendo:

- Español
- Alemán
- Inglés
- Portugués
- Chino
- Francés
- Ruso
- Italiano

Estos idiomas no se escogerán dentro de la aplicación si no que se utilizarán dependiendo del idioma en el que esté configurado el sistema operativo del dispositivo. Debido a la finalidad a la que está destinada la aplicación se ha querido evitar incluir opciones potencialmente infrautilizadas, para inclinar la balanza de opciones-simples del lado del último. Por tanto, una vez adquirido el teléfono, si el teléfono se ha adquirido en un territorio cuya lengua sea una de las mencionadas anteriormente, la aplicación elegirá automáticamente ese lenguaje.

De todas formas, es posible utilizar uno de esos idiomas cambiando el lenguaje de sistema operativo y de síntesis de voz, como se ve en la *figura 5.1*:

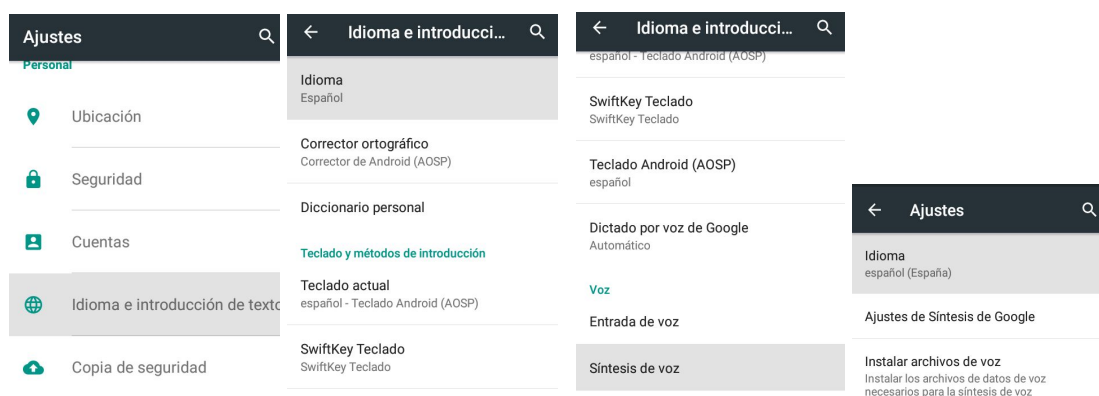


Figura 5.1 Cambiar el lenguaje del SO y de la Síntesis de voz.

Para su desarrollo simplemente se han traducido los Strings que se vayan a reproducir y se almacenan en directorios nombrados como 'values-xx' siendo 'xx' la abreviatura normalizada del idioma

elegido. Una vez realizado esto la configuración tanto de textos como de síntesis de voz estará correctamente establecida, como se ve en la *figura 5.2*.

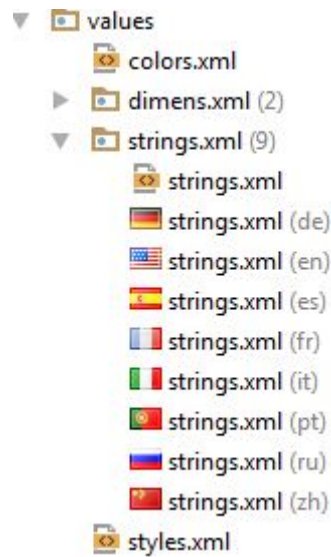


Figura 5.2 Disposición de valores de idiomas en Android Studio.

5.2 Estados

-A lo largo de la ejecución de la aplicación, está entrará en distintos estados donde se realizarán acciones determinadas y se pasará al siguiente proceso sucesivamente. La búsqueda de la imagen facial se realizará constantemente a modo de bucle, sin embargo se utilizarán distintos contadores y condiciones para controlar la frecuencia de mensajes emitidos y no suponer un elemento invasivo para el usuario. Lo que se busca con todo esto es ser complementario a las acciones del usuario y que no suponga un centro de atención constante sino un elemento pasivo.

A continuación se presentan los distintos estados por los que pasa el proceso de ejecución.

5.2.1 Localización

En el estado de Localización se realiza una búsqueda de la imagen facial utilizando el stream de imágenes proporcionado por la cámara del dispositivo.

No habrá interacción con el usuario para comenzar la ejecución de este (ni de ningún) proceso o estado, sino que estos se iterarán automáticamente.

Una vez iniciada la aplicación se ejecutará el proceso de búsqueda continuamente, dependiendo de si se obtiene un resultado positivo o negativo se realizarán acciones pertinentes. También se llevará a cabo un sistema de conteo de iteraciones para dilucidar cuánto tiempo pasa entre obtener un positivo o un negativo. El motivo de esto se explicará más adelante.

Si no se encuentra una imagen facial en el flujo de imágenes de la cámara del teléfono, se procederá a realizar una espera no sin antes cambiar las variables booleanas referidas al resultado obtenido en la iteración anterior. Estas variables nos servirán para saber si se deberá emitir un mensaje de “Se ha

perdido la cara” o comenzar con la descripción de apariencia de la cara encontrada. Una vez encontrada la imagen facial y comprobado el valor de la iteración anterior, se reservará el valor del tiempo actual en el que se da por perdida la cara.

Se comparará entonces el tiempo en el que se da por perdida la cara con el tiempo donde se encontró una cara. Esto lo utilizaremos a modo de filtro paso bajo en pérdidas breves de la localización de la cara. Se ha comprobado que resulta invasivo el hecho de que se emitan nuevos mensajes de apariencia o localización facial cuando ocurre una pérdida de la imagen facial de manera breve. Con esto evitamos ese problema ignorando las pequeñas pérdidas debido a cambio en la luz o movimientos de la cámara o la cara.

Cabe mencionar que el tiempo se medirá en número de iteraciones y no en valor absoluto en milisegundos. Esto será un aspecto a mejorar dado que dependerá de los tiempos de ejecución y potencia del procesador del teléfono, sin embargo es un sistema robusto ya que existe una dependencia en este aspecto con todos los demás procesos, lo que no llevará a desequilibrios en la duración de cada proceso, sino que simplemente, si los tiempos de ejecución del procesado de iteraciones es más rápido, los tiempos serán equitativamente más rápidos.

Este proceso se muestra en el código de la *figura 5.3*:

```
//Si no encuentra cara
if (list.size() == 0) {
    //cfind indica si es posible pasar de no encontrar cara a encontrarla
    cfind = true;
    if(cmiss) {
        //actualización tiempos de pérdida y de encontrar cara
        tmiss = tcara;
        cmiss = false;
    }
    contadorAp = 0;

    //Si ha pasado un tiempo desde que ha perdido la cara
    if((tcara-tmiss)>30 && bfind){
        bfind = false;
        bmiss = true; //
        speech("Se ha perdido la cara.");
        logoactivo.setImageResource(R.drawable.logorojo);
        dichoAp = false;
    }
}
```

Figura 5.3 Código relativo a la no-localización de imagen facial.

5.2.2 Descripción de apariencia

Uno de los primeros elementos a tomar en cuenta en la toma de características de la imagen facial es la descripción de apariencia. En este estado se proporcionará al usuario información relacionada con el género de la persona de la imagen, así como la edad y la etnia a la que más se asemeja.

Para llevar a cabo este proceso debemos de asegurar haber acabado el proceso anterior de Localización, esto conlleva no sólo haber encontrado una imagen facial sino respetar los tiempos pertinentes para ignorar aquellas pérdidas locales o breves de la cara.

Una vez asegurada la espera nos disponemos a iterar sucesivamente los procesos de detección facial referidos a la apariencia. En nuestro caso en primer lugar se tomará información del género. Se creará previamente una variable de tipo String, donde se albergará la información de la apariencia; por tanto una vez obtenida la información sobre el género se reservará en dicha variable, llamada esta vez ‘Apstr’, como se ve en la *figura 5.5*. A continuación se irá obteniendo la información relativa a los aspectos restantes referidas a la apariencia: edad y etnia. El detector de edad ofrecerá el rango de edad en el que se encuentra la persona. El detector de etnia ofrecerá un pronóstico sobre la etnia de la persona. Nótese que esta característica suele ser menos acertada en personas procedentes de Asia o África. Y se ha optado por incluir en el String a reproducir por la salida de audio la frase “parece de”.

```
public void parlanteAppearance(Face face){  
  
    switch (face.appearance.getGender()) {  
        case UNKNOWN:  
            break;  
        case MALE:  
            Apstr= getString(R.string.hombre);  
            contadorAp++;  
            //speech();  
            break;  
        case FEMALE:  
            Apstr = getString(R.string.mujer);  
            contadorAp++;  
            break;  
    }  
  
    switch (face.appearance.getAge()) {  
        case AGE_UNKNOWN:  
            break;  
        case AGE_UNDER_18:  
            Apstr = Apstr + getString(R.string.edad_sub18);  
            break;  
        case AGE_18_24:  
            Apstr = Apstr + getString(R.string.edad_18_24);  
            break;  
        case AGE_25_34:  
            Apstr = Apstr + getString(R.string.edad_25_34);  
            break;  
        case AGE_35_44:  
            Apstr = Apstr + getString(R.string.edad_35_44);  
            break;  
        case AGE_45_54:  
            Apstr = Apstr + getString(R.string.edad_45_54);  
            break;  
        case AGE_55_64:  
            Apstr = Apstr + getString(R.string.edad_55_64);  
            break;  
        case AGE_65_PLUS:  
            Apstr = Apstr + getString(R.string.edad_sobre_64);  
            break;  
    }  
}
```

Figura 5.5 Localización y reproducción por salida de audio de la apariencia.

Se debe mencionar que, una vez más, esta y todas las frases compuestas por variables de tipo String que se vayan a reproducir por la salida de audio deben ser referenciadas al archivo de formato xml referido para cada directorio de idioma ‘values-xx’ siendo xx la abreviatura del idioma pertinente, como se ha explicado en el apartado anterior.

A modo de ejemplo, el mensaje reproducido una vez localizada la cara y obtenido la información referida a la apariencia será:

“Es una mujer. De entre 35 a 44 años. Y parece de etnia del sudeste asiático”

Este tipo de mensaje referido a la apariencia no se volverá a reproducir, a no ser que se confirme la pérdida de la cara. En ese caso se repetirá el proceso.

5.2.3 Descriptor de emociones

Una vez reproducida por la salida de audio la apariencia de la cara se procederá analizar el estado emocional en el que se encuentra la persona utilizando la expresión facial.

Las emociones con las que se trabajarán serán:

- Tristeza
- Enfado
- Alegría
- Desprecio
- Sorpresa

Será muy probable que la expresión facial de una de estas emociones se prolongue en el tiempo, por lo tanto es posible localizar en cada iteración la aparición de dicha expresión. Sin embargo sería indeseable, ya que sería muy invasivo, emitir un mensaje en cada ejecución enunciando la emoción de la expresión facial.

Es por ello que se realizan operaciones con variables booleanas para discernir si el mensaje ya ha sido emitido (*figura 5.7*), así como variables para contar si esta emoción ha perdurado lo suficiente en el tiempo como para considerarse una emoción y no un evento casual, como se ve en la *figura 5.6*. Por ejemplo, es posible que una persona reaccione con una expresión de tristeza ante un estímulo, lo cual se considerará un evento de la interacción. Pero también es posible que el estado emocional de esa persona a lo largo de una interacción sea de tristeza, y se manifestará con una expresión facial de tristeza prolongada en el tiempo.

```
public void parlanteEmotion(Face face){  
  
    if(face.emotions.getJoy()>50){  
        //si no ha cambiado de emocion  
        if(emocion==1) {  
            contador++;  
        }  
        else {  
            contador=0; //se resetea su con  
            dicho=false;  
            emocion=1;  
        }  
    }  
    else if(face.emotions.getAnger()>50){
```

Figura 5.6 Detección de emoción y uso del contador.

```

if(contador>3) {
    switch (emocion){
        case 1:
            if(!dicho) {
                speech("Está alegre");
                dicho=true;
                logoactivo.setImageResource(R.drawable.logofeliz);
                contador=5; //limitador contador
            }
            break;
        case 2:

```

Figura 5.7 Reproducción por salida de audio después de la espera.

A modo de ejemplo un mensaje que se reproducirá por la salida de audio si nos encontramos ante una expresión facial que evoque la emoción de tristeza será:

“Está triste.”

5.2.4 Eventos

Aquellos aspectos referidos a expresiones relativamente cortas en el tiempo, o microexpresiones como cerrar los ojos, bostezar, levantar las cejas, etcétera; se considerarán eventos.

La naturaleza de los eventos es distinta a la de los elementos anteriores, ya que estos ocurren desde el inicio de la interacción, de manera continuada o intermitente. Por lo tanto estas expresiones hay que tratarlas de modo distinto en relación a la periodicidad de reproducción de mensajes.

Es posible que los eventos ocurran en un espacio temporal muy corto, sin embargo se ha optado por utilizar una ventana de tiempo desde que ocurre un evento y se emite un mensaje por la salida de audio hasta que se pueda emitir un mensaje nuevo referido a un nuevo evento. Esta decisión de diseño se toma para que la aplicación no reproduzca mensajes de manera muy continuada de forma que llegue a resultar molesto al usuario.

Existen un notable número de expresiones que pueden entrar dentro de los eventos, sin embargo en esta versión se han utilizado las siguientes:

- Cerrar los ojos.
- Las expresiones similares a los emojis descritos en el apartado 2.4.

El motivo por el cual se han elegido estas expresiones es la robustez de las mismas. Se han realizado pruebas también con otras expresiones, como por ejemplo ‘abrir la boca’ para relacionarlo a su vez con un bostezo, o morderse los labios; sin embargo nos encontrábamos con un gran número de falsos positivos, lo que conllevaba a un gran número de mensajes reproducidos y a su vez una ligera molestia en el uso de la herramienta. A medida que los procedimientos de reconocimiento de expresiones mejore se podrán ir agregando en este campo aquellos gestos que se consideren oportunos, y así conseguir un mayor abanico de expresiones con el que operar.

5.3 Interfaz

-La interfaz de la aplicación es simple y sencilla. Debido a la funcionalidad de la herramienta, el usuario al que va destinado no podrá disfrutar de la interfaz visual, sin embargo se ha optado por transmitir toda la información que la salida de audio emite de forma visual y sencilla.

Al iniciar la aplicación aparecerá un splash (pantalla de carga) con el logotipo de la misma. Una vez arranque podremos observar 3 elementos en la pantalla:

- Feedback del stream de la cámara
- Botón para utilizar a la cámara frontal o trasera
- Icono que proporcionará la información reproducida por la salida de audio

Feedback de cámara:

-Sin realizar ninguna interacción con la aplicación, automáticamente aparecerá el feedback de la cámara con la búsqueda de información facial activada.

Se respetarán las relaciones de aspecto de la pantalla de cada dispositivo, sin embargo, dado que cada teléfono posee características distintas también de la resolución de la cámara ésta se verá modificada pertinentemente para encajar dicha imagen en el display, respetando la relación de aspecto y de esta manera no deformando la imagen obtenida.


Botón frontal/trasera:






-Este botón será accesible para su localización en Google TalkBack y permitirá al usuario alternar en el uso de la cámara trasera o delantera del teléfono. Esto permitirá adaptarse a la situación y modo de empleo en el que el usuario utilice la herramienta.

Iconos de información:

-Toda la información que se transmita por la salida de audio será posible de obtener también por medio visual mediante iconos informativos que se irán cambiando a medida que ocurran cambios en las expresiones faciales.

En la *tabla 5.1* se muestran los iconos utilizados y la situación asociada que debe ocurrir para que se muestren:

CARA NO DETECTADA	
-------------------	---

<p>CARA DETECTADA</p> <p>EMOCIÓN: NEUTRAL</p>	
<p>EMOCIÓN: ALEGRE</p> <p>EVENTO: SONREÍR</p>	
<p>EMOCIÓN: TRISTEZA</p> <p>EVENTO: ENTRISTECERSE</p>	
<p>EMOCIÓN: ENFADO</p> <p>EVENTO: ENFADARSE</p>	
<p>EMOCIÓN: SORPRENDIDO</p> <p>EVENTO: SORPRENDERSE</p>	

<p>EVENTO: MANDAR UN BESO</p>	
<p>EVENTO: SACAR LA LENGUA</p>	
<p>EVENTO: SACAR LA LENGUA Y GUIÑAR UN OJO</p>	
<p>EMOCIÓN: DESDÉN EVENTO: SONRISA DESPECTIVA</p>	

<p>EVENTO: SONRISA MOSTRANDO LOS DIENTES</p>	
<p>EVENTO: CERRAR LOS OJOS</p>	

Tabla 5.1 Iconos designados para emociones y eventos.

Capítulo 6

EVALUACIÓN DEL SISTEMA

6.1 Resultados de la evaluación

-A continuación se mostrarán los resultados del cuestionario de evaluación utilizado para obtener feedback de los usuarios sobre distintos aspectos.

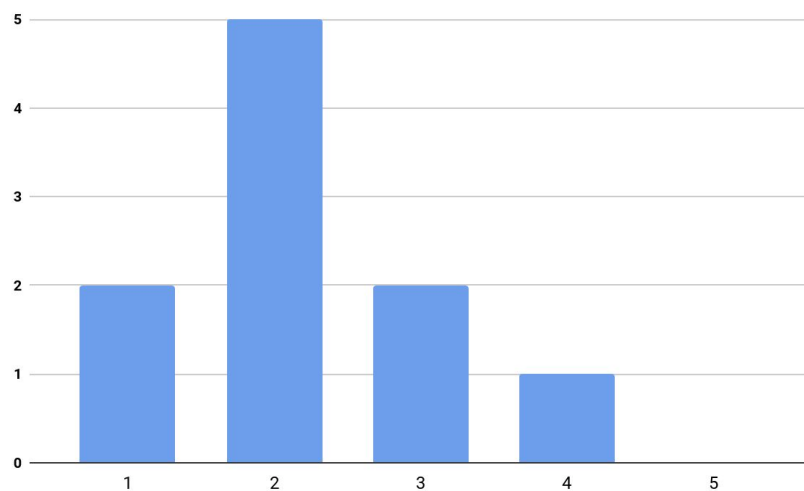
Las preguntas formuladas se centrarán en distintos aspectos importantes sobre la funcionalidad de la aplicación, tales como facilidad de uso, opinión sobre su eficacia, invasividad, etcétera.

Desafortunadamente no se ha podido consultar a personas con discapacidad visual, sin embargo los usuarios que han utilizado la aplicación lo han hecho sin contar con visibilidad. Por tanto no se han visto beneficiados de la visión para su uso.

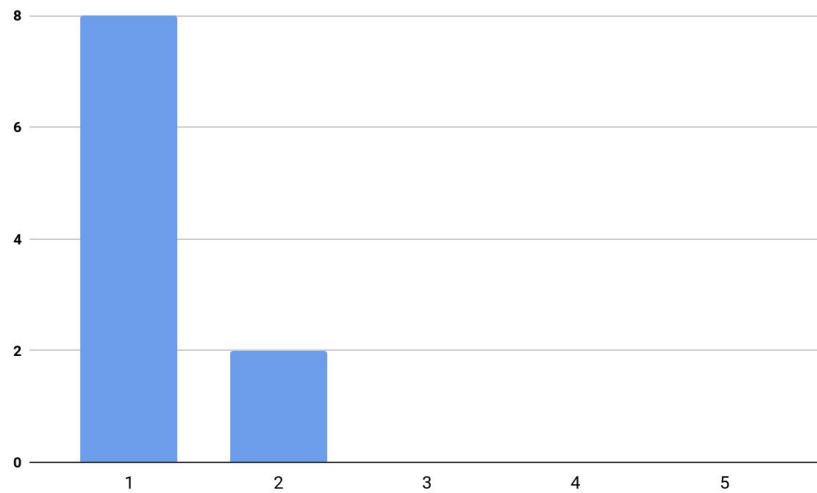
Posteriormente se ha pedido utilizar la aplicación de manera normal para observar la interfaz gráfica.

Las preguntas se han realizado a un total de 10 usuarios. Seguidamente se muestran los resultados:

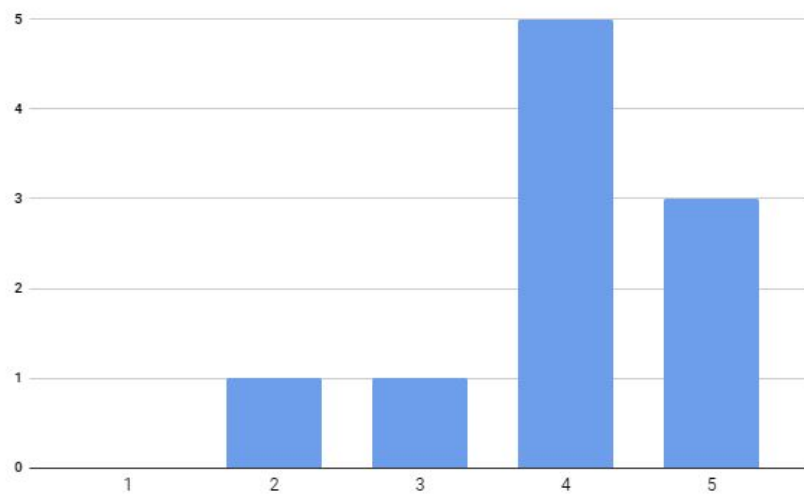
- En una escala del 1 al 5 ¿qué grado de dificultad posee el manejo de la aplicación con la visibilidad impedida? (siendo 5 máxima dificultad y 1 la mínima)



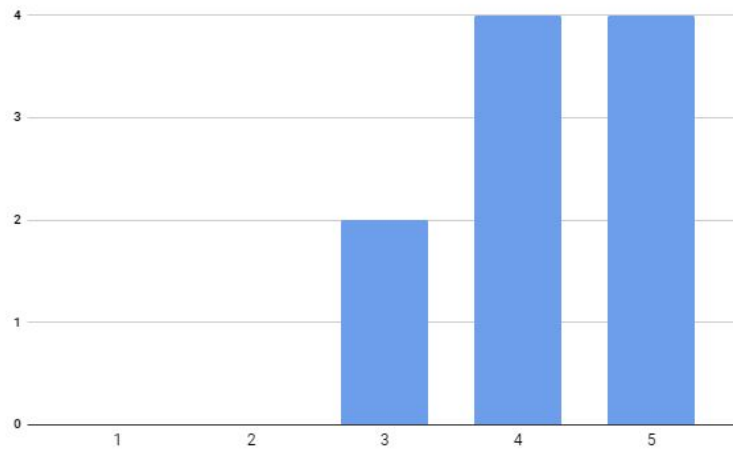
- En una escala del 1 al 5 ¿qué grado de dificultad posee el manejo de la aplicación utilizando la visión? (siendo 5 máxima dificultad y 1 la mínima)



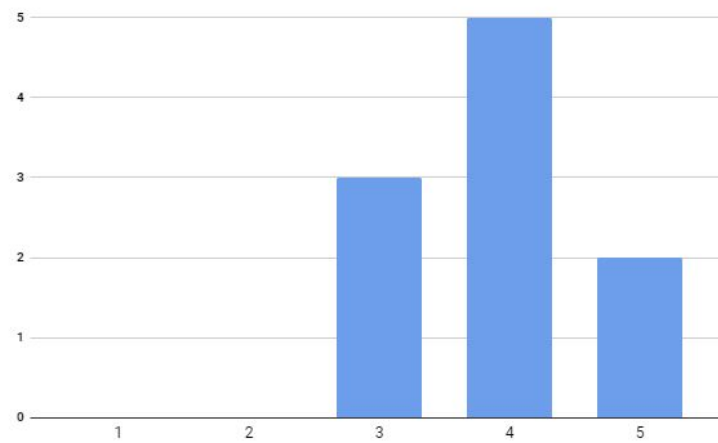
- En una escala del 1 al 5 ¿qué grado de molestia le produce la aplicación durante una interacción? (siendo 5 máxima molesta y 1 mínima molestia)



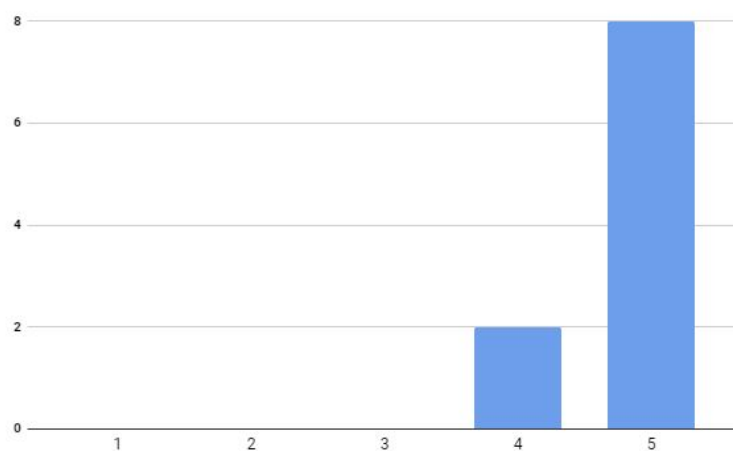
- En una escala del 1 al 5 ¿qué grado de afinidad ocurre entre las frases reproducidas y su interlocutor? (siendo 5 máxima afinidad y 1 mínima afinidad)



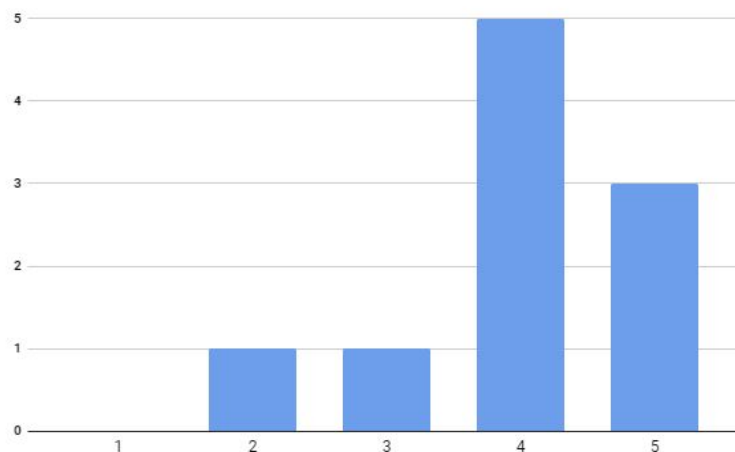
- En una escala del 1 al 5 ¿se encuentra cómodo con la interfaz visual y los iconos utilizados? (siendo 5 muy de acuerdo y 1 muy en desacuerdo)



- En una escala del 1 al 5 ¿cree que la aplicación es útil para personas con ceguera? (siendo 5 máxima utilidad y 1 mínima utilidad)



En una escala del 1 al 5 ¿cree que son pertinentes las emociones, apariencia y eventos utilizados? (siendo 5 muy de acuerdo y 1 muy en desacuerdo)



6.2 Conclusiones de la evaluación

-Una vez obtenidos los resultados de la evaluación se pueden sacar algunas conclusiones sobre el manejo y eficacia de la herramienta.

- En primer lugar, en relación a la dificultad en el uso se comprueba que es notablemente más difícil utilizar la aplicación con la capacidad visual impedida. Esto podría ser solventado con un sistema de tutorial inicial emitido por voz, sin embargo habría que considerar la opción de activar o desactivarlo para no repetir el mensaje siempre que se abra la aplicación. Sin embargo, cabe recordar que todas las opciones o interacciones a través de la interfaz táctil se volverá tedioso con la utilización de herramienta de accesibilidad como Google Talk. Por tanto se podría explorar la opción de controlar la aplicación mediante mensajes de voz.
- En segundo lugar también es considerable la molestia producida por la aplicación manteniendo una conversación. Sin embargo esto puede ser debido a la inexperiencia o poca costumbre en el manejo de la aplicación. Pero no por ello hay que considerar la posibilidad de reducir los mensajes o los estímulos que activan la emisión de mensajes por la salida de audio.
- Con respecto a la información ofrecida la opinión de los usuarios opina que es pertinente en términos generales las expresiones emocionales y apariencias ofrecidas. Y también, se cree pertinente la capacidad de ofrecer esta información a personas invidentes.
- La interfaz visual resulta relativamente cómoda y los iconos ofrecen información simple y pertinente. Sin embargo se ha mencionado en algún encuestado la simplicidad de la interfaz en cuanto a su diseño.

Capítulo 7

CONCLUSIONES Y TRABAJO FUTURO

7.1 Conclusiones

-Una vez finalizado el proyecto debemos repasar aquellos aspectos que se han logrado cumplir y también aquellos que necesitan mejoras o simplemente no se han conseguido.

En primer lugar aspirábamos a construir una herramienta que permita proporcionar un nuevo canal comunicativo a aquellas personas que sufren una discapacidad visual que les impide beneficiarse de una comunicación corporal manifestada visualmente. Esto se ha logrado y pone a disposición del usuario discapacitado un nuevo plano de comunicación, que, aunque sea de información algo reducida, otorga un método de interacción irrealizable hasta el momento. Esto da como fruto una posible nueva forma de comunicarse ya que se podrá tomar en cuenta campos como el aspecto o expresiones faciales en las interacciones llevadas a cabo por el usuario, lo que moverá al usuario a acercarse a contextos y situaciones en las que esto sea más importante.

En segundo lugar la característica pasiva o complementaria a las acciones cotidianas del usuario se ha logrado relativamente. La herramienta resulta relativamente invasiva al emitir mensajes relativos a eventos o pérdidas en la localización de la cara del interlocutor, esto se debe mayoritariamente a los algoritmos de localización facial y de expresiones. Sin embargo, este aspecto mejorará considerablemente en escenarios que cumplan con solvencia los requisitos ambientales referidos a la luz o posición de la cara. Además, la actualización de los algoritmos de búsqueda y utilización de los recursos disponibles es constante. Se disponen de nuevas actualizaciones en este aspecto con bastante frecuencia lo que permite pronosticar una mejora de las prestaciones en este aspecto a medio y largo plazo.

El tercer lugar corresponde a la accesibilidad que la aplicación posee para ser adquirida y utilizada por usuarios invidentes. Este aspecto se resuelve con relativa solvencia ya que, pese a obtener versiones para la mayoría del parque de móviles mundial, no existe versión para iOS. Esto podría ser solucionable adquiriendo la licencia correspondiente ya que existe soporte para el sistema operativo mencionado. Además los recursos técnicos utilizados no supone una barrera ya que estos no son altos en el panorama actual de móviles del mercado.

En el ámbito personal este proyecto me ha servido mucho en diferentes campos.

He conocido las tecnologías apoyadas en el análisis de sentimientos y reconocimiento de emociones en sus distintas variantes: voz, texto e imagen. Me doy cuenta ahora de lo mucho que se está avanzando en este campo cuando veo a teléfonos móviles o sistemas de marketing apoyarse en ellos para automatizar estrategias de mercado o para mejorar los sistemas de seguridad.

También he aprendido sobre las tecnologías de accesibilidad disponibles para todo tipo de discapacidades, no solo visuales sino también auditivas o cognitivas. Poco a poco aprendo a ver más y más relaciones entre las discapacidades y la tecnología, sobre todo las de las últimas décadas, que proporciona una mayor facilidad para aprender conceptos o llevar a cabo interacciones con el entorno. Es un campo muy importante para personas que sufren estas dolencias y a estudiar y desarrollar en las próximas décadas, con la importancia que conlleva mejorar la calidad de vida de este colectivo.

En definitiva he aprendido mucho sobre la tecnología, la discapacidad, y también sobre conocimientos técnicos en problemas que he tenido que resolver en el desarrollo del proyecto. Todo ello me invita a seguir y tener de aquí en adelante un ojo pendiente de cómo evoluciona el campo de las tecnologías destinadas a personas con discapacidad.

7.2 Trabajo futuro

-A lo largo del desarrollo de la aplicación móvil han surgido nuevos elementos no planteados al inicio del proceso. Algunos han sido implementados como por ejemplo el uso de diferentes iconos para representar los diferentes estados emocionales o los eventos ocurridos en la interlocución. Sin embargo existen aspectos que no se han implementado y que podrían suponer una mejora de características. A continuación se muestra una lista de posibles mejoras para el futuro de la herramienta:

En primer lugar es de vital importancia para la funcionalidad de la aplicación el pulido y mejora de la invasividad de la aplicación. Se considera este aspecto muy importante ya que es lo que más afecta a la experiencia del usuario. Si idealmente el usuario posee un auricular conectado a la salida de audio mientras mantiene un diálogo con un interlocutor, la intermitencia indeseada de los mensajes va afectar negativamente al usuario. Es por tanto que la mejora de este aspecto resulta crucial.

Hasta ahora las esperas utilizadas para emitir mensajes simplemente se basan en esperas fijas de tiempo, sin embargo esto podría mejorarse realizando una dependencia entre los tiempos de espera y el diálogo en sí del usuario. Sería una mejora notable, por ejemplo, tomar en cuenta las pausas que ocurren en la conversación, utilizando para ello el micrófono del dispositivo y emitir mensajes en las etapas de silencio de los interlocutores.

Otra mejora en este aspecto podría ser implementar la capacidad de poder controlar la cascada de mensajes emitidos mediante control por voz. Utilizando palabras clave que puedan ser reconocidas por la aplicación y consiguiendo de esta forma que la aplicación solo entre en juego cuando el usuario, mediante la voz, lo considere oportuno. Como se puede observar cualquier solución debe tomar en consideración las discapacidades del usuario y la no invasividad en su acción concurrente.

En segundo lugar se podría ampliar el campo de las funcionalidades de la aplicación implementando un sistema de reconocimiento facial que permita diferenciar e identificar imágenes faciales. Esta aplicación podría ser útil en situaciones en las que el usuario deba tratar con un gran número de personas, de tal forma que sea incapaz de reconocer a una persona mediante el tono y timbre de su voz. A su vez se puede implementar un sistema de contactos que funcione utilizando los datos ofrecidos por la identificación. Es decir, guardando una imagen facial a modo información de contacto

en la libreta de contactos y pudiendo sustraer datos de la misma en una futura situación en la que la aplicación analice esa misma cara.

En tercer lugar se podrían incluir características más sencillas de implementar pero que pueden mejorar las prestaciones de la aplicación ya diseñada. Por ejemplo el uso de la linterna del dispositivo para poder llegar a esas condiciones de iluminación necesaria en caso de utilizarse en un ambiente de baja visibilidad. O también un diseño de interfaz gráfico más atractivo y que ofrezca más opciones para una versión de la aplicación más avanzada o personalizable.

Aún y todo cabe mencionar que las herramientas utilizadas para la identificación de aspecto y emociones está en continuo proceso de mejora. Afectiva ofrece actualizaciones regularmente mejorando las prestaciones y aportando nuevas características disponibles. Para estar informado de estas y otras noticias relacionadas con la herramienta y el panorama de análisis de sentimientos y reconocimiento de emociones se sugiere estar al tanto de sus redes sociales o página web.

Capítulo 8

GESTIÓN DEL PROYECTO

8.1 Planificación temporal

-La planificación del proyecto está desglosada en los apartados adjuntos en el ANEXO A, presentado y definido en el capítulo **1.3 Fases de desarrollo**. En la *figura 8.1* se muestra el diagrama de Gantt basado en dichas fases de elaboración del proyecto.

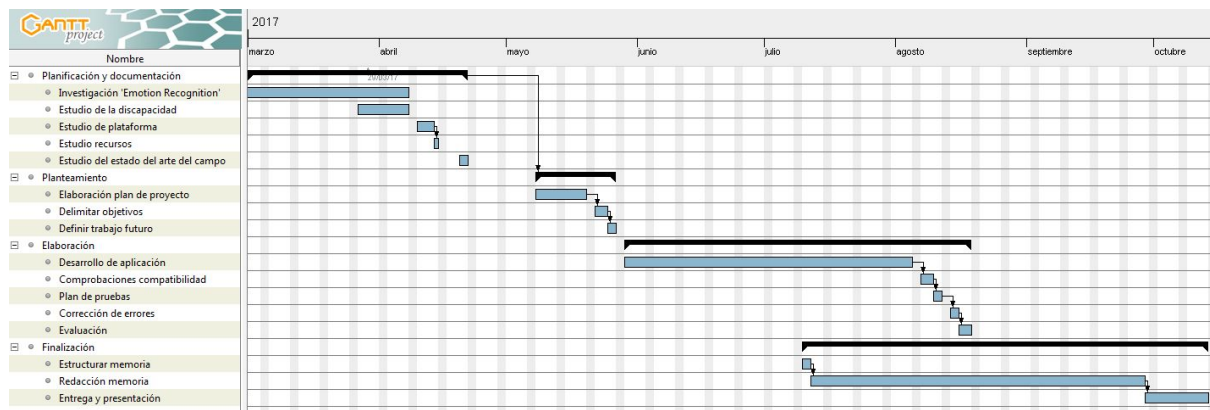


Figura 8.1 Diagrama de Gantt del proyecto.

En el mismo se comprueba el calendario de tareas realizado y la sucesión correspondiente a las tareas, aplicando en cada caso si lo requiriera las tareas antecesoras de otras. Se muestra también que la elaboración de la aplicación y la redacción de la memoria son las tareas más largas, y además, que la última se empieza a realizar pasado el ecuador de la fase de elaboración.

El proyecto se comienza en la semana 9 y se finaliza en la semana 42. A una media de 3 horas diarias dedicadas a la realización del proyecto supondría un total de 495 horas.

En la *figura 8.2* se muestra el calendario previsto de las tareas realizadas desglosadas.



Nombre	Fecha de ini...	Fecha de fin
☐ • Planificación y documentación	27/02/17	21/04/17
• Investigación 'Emotion Recognition'	27/02/17	7/04/17
• Estudio de la discapacidad	27/03/17	7/04/17
• Estudio de plataforma	10/04/17	13/04/17
• Estudio recursos	14/04/17	14/04/17
• Estudio del estado del arte del campo	20/04/17	21/04/17
☐ • Planteamiento	8/05/17	26/05/17
• Elaboración plan de proyecto	8/05/17	19/05/17
• Delimitar objetivos	22/05/17	24/05/17
• Definir trabajo futuro	25/05/17	26/05/17
☐ • Elaboración	29/05/17	18/08/17
• Desarrollo de aplicación	29/05/17	4/08/17
• Comprobaciones compatibilidad	7/08/17	9/08/17
• Plan de pruebas	10/08/17	11/08/17
• Corrección de errores	14/08/17	15/08/17
• Evaluación	16/08/17	18/08/17
☐ • Finalización	10/07/17	13/10/17
• Estructurar memoria	10/07/17	11/07/17
• Redacción memoria	12/07/17	28/09/17
• Entrega y presentación	29/09/17	13/10/17

Figura 8.2 Desglose de las tareas del proyecto.

8.2 Presupuesto

-Se incluye a continuación el desglose presupuestario del proyecto, incluyendo costes directos, relativo a los costes de personal y de equipo, y lo costes indirectos, siendo este el 20% del monto total.

Costes de personal

La siguiente fórmula calculará el coste directo de personal:

$$\text{Coste} = ((\text{días} * \text{horas diarias}) / \text{horas dedicadas empleado por mes}) * \text{coste empleado mes}$$

Duración: 33 semanas=165 días laborables

Horas diarias: 3 horas

Horas dedicadas empleado por mes= 60 horas

Coste empleado al mes = 2250 €

Coste total= 18562.50 €

Coste de equipo

A continuación se desglosan los costes relativos al equipo, tanto de hardware como software:

- **Hardware**

- Ordenador: 800 €
- Dispositivo móvil ZTE Blade L5: 60 €
- Cable usb-microusb: 3 €

- **Software**

- Android Studio: 0 €
- SDK version 4.4 y 5.0 de Android: 0 €
- JDK: 0 €
- Photoshop licencia 9 meses: 324 €
- Google Drive: 0 €
- Licencia gratuita de Afectiva: 0 €

La *tabla 8.1* muestra la amortización de los equipos utilizados, en él se muestra el coste imputable de cada equipo siendo este calculado como:

$$(A/B)*C*D$$

Donde:

A es el número de meses desde la facturación del equipo

B es el periodo de depreciación

C es el coste del equipo sin IVA

D es el porcentaje de uso aplicado al proyecto

Equipo	Coste sin IVA	% Uso dedicado al proyecto	Dedicación (meses)	Periodo de depreciación	Coste imputable
Ordenador	632	100	8.25	60	86.9
ZTE Blade L5	47.4	100	8.25	60	6.51
Photoshop	255.96	100	8.25	60	36.26
Cable usb-microusb	2.37	100	8.25	60	0.33
TOTAL					130 €

Tabla 8.1: Amortización de las herramientas utilizadas.

El precio final del proyecto es calculado sumando los costes directos e indirectos y el IVA correspondiente, como se muestra en la *tabla 8.2*.

Descripción	Coste
Costes de personal	18562.50 €
Costes de equipo	130 €
Costes indirectos	3738.50 €
TOTAL SIN IVA	22431 €
IVA (21%)	4710.51 €
TOTAL	27141.51 €

Tabla 8.2: Costes totales del proyecto.

8.3 Impacto socio-económico

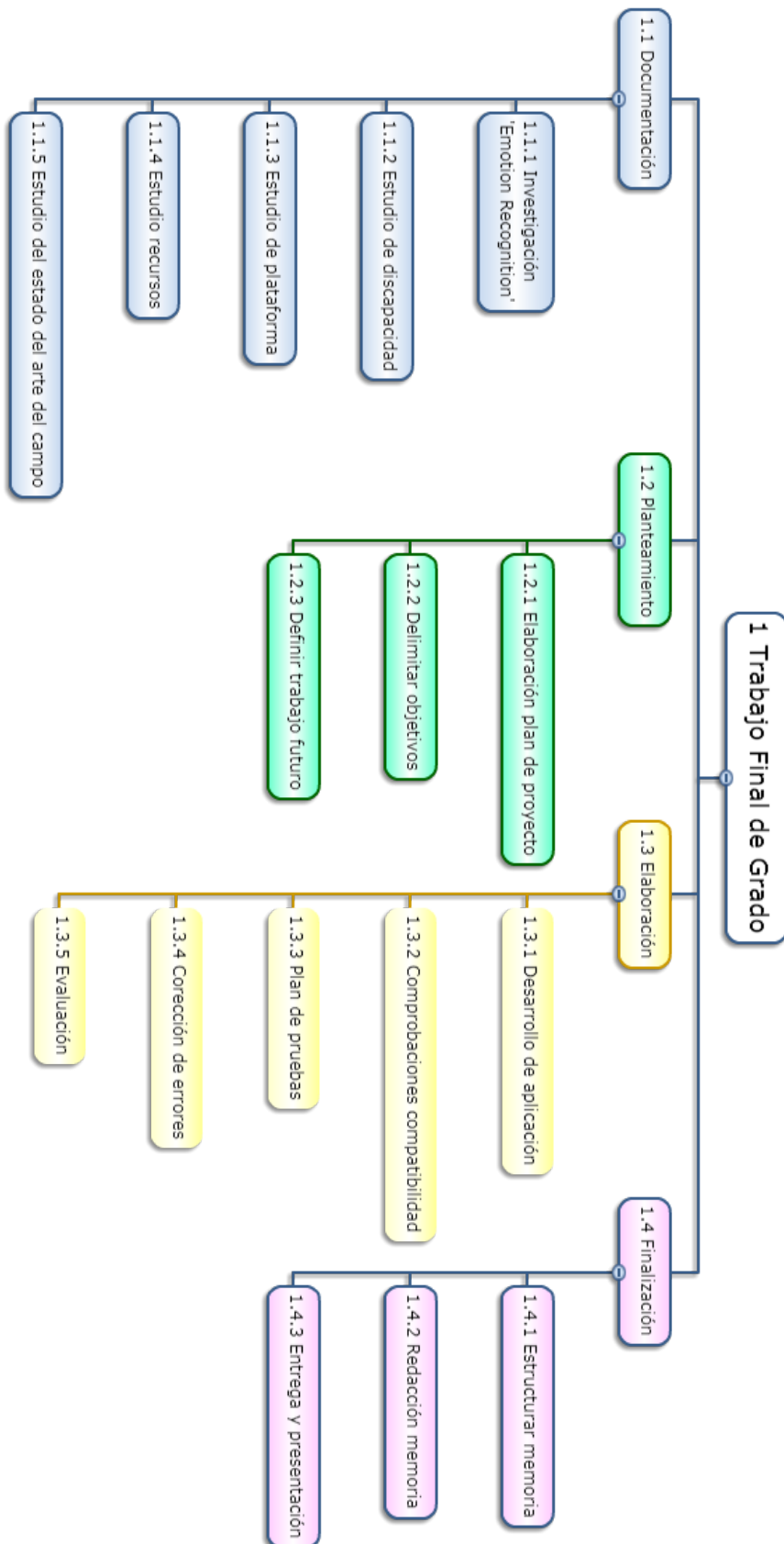
-Hay que mencionar también el impacto que podría suponer esta aplicación en la sociedad. La implicación que puede causar está a la par de cualquier avance tecnológico aplicado a conjuntos de la sociedad con discapacidad.

En esta aplicación se presenta unos beneficios para personas invidentes inmediatos. Ofrece la oportunidad de disfrutar de un canal comunicativo privado para este grupo de la sociedad, que es la comunicación corporal y las comunicación por medio de las expresiones faciales. Esto puede influir a las personas en aspectos positivos tales como ofrecerles mayor grado de independencia para, por ejemplo, identificar personas o mejorar sus interacciones verbales incluyendo un grado de comunicación facial expresiva.

La importancia de mejorar la calidad de vida de las personas con discapacidad y en este caso de discapacidad visual es enorme. No solo contribuye positivamente a la sociedad sino que también puede ofrecer oportunidades laborales nuevas para esta minoría. Escenarios laborales donde se requiera identificación de personas o aquellos ámbitos donde la interacción cara a cara sea predominante podrán ser entonces los nuevos escenarios de personas con discapacidad visual. Si se sintetizan los escenarios visuales, con la ayuda de la tecnología de reconocimiento de emociones poco a poco será posible desempeñar las mismas tareas que una persona con visión.

Todo esto aplicado, como se ha visto en el capítulo 1 de la memoria, a las regiones afectadas por discapacidades visuales, y tomando en consideración los ocho idiomas en los que la aplicación se ha traducido puede producir un aumento considerable de la economía y empleo, teniendo en cuenta el tipo de usuario final al que está enfocada la herramienta.

ANEXO A: Fases de desarrollo



ANEXO B: English content

-Along with the exponential advance on computational field, information processing and telematics of the last decades, our daily life has changed markedly. It has been introduced on our day to day new ways of communication, more facilities for communicating through thousand kilometers of distance, instantaneity of our ideas... We have witnessed how, in the period of time of our existence, we have gone from sending letters to another continent without knowing if they will arrive, to share a breakfast with a relative by videoconference, free of charge, using our mobile phone.

However, what about the person who, before the technological boom, had difficulty communicating more basic ideas than all that? A smile, a kiss, a tone of voice, a fuss, are just some examples of messages that are trivial for the most of us and we do use constantly in our day to day, but for many people with disabilities it is a desire to experience it .

Will these technological advances have been able to solve this?

Of all the disabilities that affect the communication of basic ideas we are going to focus on the **blindness**.

We will study the context of this disability at a national and global level to obtain clues that can help us to face a design and development of a mobile application.

- *National scope*

i) Data from the National Institute of Statistics

Figure B.1 shows the number of people with disabilities who are unable to perceive any image. This upward trend is evident from the age of 65, with 73% of the total ranging from this age onwards.

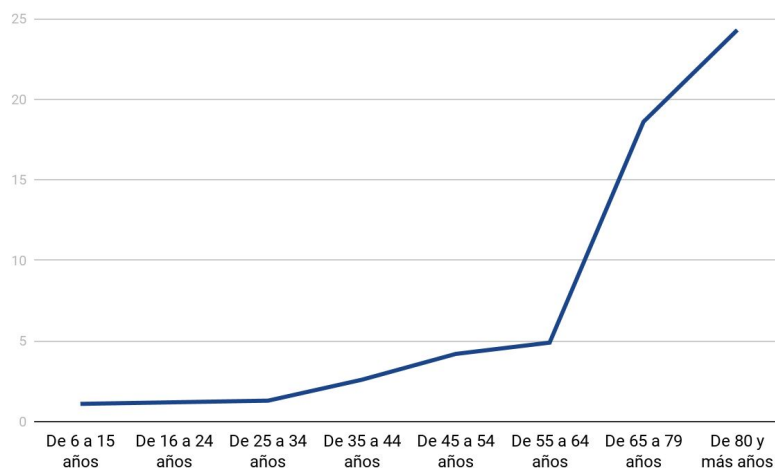


Figure B.1 Disabled (in thousands) to perceive any image according to age

ii) Data of members of ONCE

The ONCE has a total of 72,097 members in June 2017.

Figure B.2 distributes it according to degree of disability:

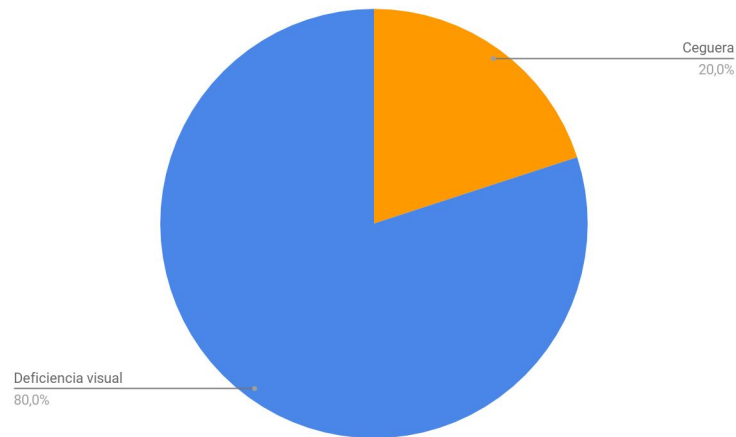


Figure B.2 Proportion of blindness and low vision (millions).

Defining 'blindness' (ceguera) as absence of vision or only perception of light, and 'visual impairment' (deficiencia visual) as the ability to maintain a functional visual rest for daily life.

In short, 14,419 members of ONCE suffer total blindness.

- *International scope*

In this context, the World Health Organization (WHO) is preparing a global plan of action for 2014-2019 on universal eye health. It will offer the following data for 2010:

4.1% of the world's population has some kind of visual disability, equivalent to 285 million people.

246 million people have low vision problems, and the remaining amount, 39 million, have a complete loss of vision.

The most affected countries are from the Africa, Arabia, South and Southeast Asia and Oceania. This gives us a clue about what languages we have to use in case of wanting to enlarge the range of potential users of the application.

With this information we can devise and plan a series of objectives that provide solutions to people affected by this type of disability.

The objectives chosen are:

-Build an application that provides a type of communication otherwise inaccessible to a blind person. We will choose facial communication and expressions. The application will be able to obtain facial information from the interlocutor placed in front of the device and provide it to the blind using spoken messages.

-Given the information presented above concerning the age ranges, the application should be simple and easy to use, since the people most affected are elderly. There are standard accessibility tools in mobile operating systems that allow the ordinary use of the device, however these are tedious and difficult to use. Therefore we will try not to depend on these tools.

-The WHO report tells us in which languages the application should be translated. Finally we are choosing Italian, French, Spanish, English, German, Portuguese, Russian and Chinese.

-Due to mostly all the regions are located in developing countries, it would be optimal that the application does not require high technical hardware characteristics, taking as the bottom margin Android 4.4.

-In this and any project it is of vital importance an approach of the development phases that are going to execute. A poorly designed project around its management, development, delimitation of objectives and deadlines of compliance is destined to failure.

To do this, we must divide the problems into smaller problems and resolve them gradually.

In **ANEXO A: Fases de desarrollo** shows the Work Breakdown Structure (WBS) used to break down the project into different stages and sub-stages.

This WBS is composed of 4 main stages destined to tasks of different aspects.

- **Stage 1.1 Documentation (Documentación)**

- **1.1.1 Research in Emotion Recognition':** Documentation and research related to mechanisms of emotion recognition and sentiment analysis. Study of the types, selection and search of tools for their performance.
- **1.1.2 Disability study:** After the election, study of possible ends. It is selected an accessibility purpose for the disabled. Study of gadgets and existing accessibility tools. Test of these tools.
- **1.1.3 Platform study:** Research the current state of the Android operating system. Study of necessary software, updates and hardware requirements necessary for its operation.
- **1.1.4 Resource study:** List of necessary and available resources. Checking its correct operation and preparation for using in developer mode.

- **1.1.5 Study of the state of the art of the field:** Research of the market of applications for blind people. Study of Google Talk type accessibility tools.
- **Stage 1.2 Approach (Planteamiento)**
 - **1.2.1 Development of a project plan:** Development of a development plan.
 - **1.2.2 Delimit objectives:** Definition of objectives based on available resources and capacities.
 - **1.2.3 Defining future work:** Development of a prototype of a surplus itinerary that expands the defined objectives
- **Stage 1.3 Development (Elaboración)**
 - **1.3.1 Application development:** Elaboration of the necessary code for the desired functionality. Integration of libraries and start-up. Integration of API Afectiva and Android SDK Tools.
 - **1.3.2 Compatibility checks:** Operation check in different versions of simulated operating system in IDE.
 - **1.3.3 Test plan:** Development of a non-intensive system of test plan and execution.
 - **1.3.4 Correction of errors:** Correct the source code in those erroneous aspects discovered in the test plan.
 - **1.3.5 Evaluation:** Evaluation of the application.
- **Stage 1.4 Ending (Finalización)**
 - **1.4.1 Structure memory:** Elaboration of table of contents.
 - **1.4.2 Writing memory**
 - **1.4.3 Delivery and presentation**

Resources used

-The following defines the resources of both hardware and software used for the development of the project.

- Hardware:
 - ZTE Blade L5 mobile device
 - 8 Megapixel Camera
 - Processor with 2 cores at 1.3GHz
 - 1GB of RAM
 - Computer
 - CPU AMD Phenom II X4 960T 3.00GHz
 - 8GB of RAM
 - Peripherals
 - USB-MicroUSB Cable
- Software:
 - Android Studio (v1.1 and 2.1)
 - Android SDK Tools
 - Android SDK Platform 5.0.1 and 4.4.2
 - ARM EABI v7a System Image (AMD CPU Accelerator)
 - Affective Toolkit API

Sentiment analysis and emotion recognition

-Thanks to advances in computing speed and computation of processors of the last decades it has been possible to put algorithms in the reach of the consumer that allow us to calculate the subjective sensations of a person with respect to an input of information. It is possible to deduce the emotional state of a person from an input presented in different format (video, photograph or text), or

also to decipher computationally subjective aspects of a text, such as its positivity or negativity, if it contains irony, or its nature, deducing if it is an expository text, a critic, etcetera; as well as the topic of which it treats.

The following are examples of emotion analyzers and emotion recognizers that apply to different input formats. It will be explained what they offer and the possible immediate uses in which they are applied.

◆ **Text:**

This group includes tools like IBM Watson, Microsoft Text Analytics, Text2data, Natural Language Tool Kit (NLTK), TheSayer, among many others.

Methodology:

-There are several types of processes to make a feeling analysis based on an entry in text format. Some of them will be briefly described below:

- *Keyword localization method:*
It is an easy to implement and intuitive method since it only consists of locate words that are preset within a range of feelings (sadness, joy, anger, fear, etcetera). The process analyzes the text searching for these keywords and checks the applied emphasis, then it gives results about the emotion obtained with the highest result.
- *Lexical Affinity Method:*
This method operates at the lexical level of the word, that is, it punctuates each word based on the emotion related to its meaning. The analysis of the text consists of analyzing each word and classifying them according to the emotion it evokes independently. The problem with this method is that it does not take into account the context in which each word is placed, therefore it can give rise to inaccuracies.
- *Learning methods:*
These methods try to recognize the emotions of a text by applying a previously trained classifier, which has identified texts, phrases, words and particles within an emotion category and then applied it to a set of test texts.

Use cases:

- Monitoring the audience and listening to the opinion of society. It is possible to monitor in real time the opinions about a product after an event through their opinions on social networks.
- Personalized marketing based on the client's emotions. It allows the possibility to modify the advertising messages according to the mood of the client thus providing more empathy in the advertisement, which will help its assimilation and influence.

- Customization in automated messages. Possibility of analyzing the messages in an interaction with the client and constructing a response according to the situation.
- Automated text analysis. It provides the ability to classify and synthesize texts, for example of journalistic nature, depending on the topic they are dealing with and the positivity or negativity expressed in the opinions of the text.

❖ **Voice:**

Included in this category are applications such as 'Empath' from Smartmedical Corp., 'Vokaturi' or 'Good Vibrations'

Methodology:

-The methods used for this group of emotions recognition is independent of the content of what is said in the oral messages, i.e. the analysis will fall on the tone, frequencies, speed, and so on, and not on the meaning or language used in the discourse. The sound factors to be used are generally the tone of the voice, the speed of speech, the volume and the different events that occur in a speech, such as syllables, explosive phonemes, nasals, or silences. The process of analyzing all these factors and detecting emotion in a discourse is based on learning networks and statistical classifiers that are trained and subsequently tested.

Use cases:

- Monitoring and control of feedback obtained during telephone calls in customer services. It is possible to quantify and manage information related to the satisfaction of client support or promotional support provided to a customer through voice service.
- Personalization of messages in the interaction between human-machine, thanks to the information obtained from the emotional features included in the speech.

❖ **Facial image:**

This is the group that the API we will use in this project, Affectiva, works on but there are other applications such as Emotion API from Microsoft, EmoVu, Kairos or FaceReader from Noldus, among others.

Methodology:

-Facial expressions give a lot of information about the emotions a person feels. Elements such as eyebrows, eyelids, mouth and cheeks give us the opportunity to know someone's mood without having a verbal conversation. The computational methods to obtain this information are based on the spatial distribution of certain facial points of interest.

In a first stage, these points are located using image processing techniques and geometric models. Subsequently it is tracked how these points are shifted in the image, as shown in *Figure B.3*, to take these data and apply them to a system of classifiers that will give us the probability of the most similar emotion. This is aided by simple facial movements, such as an eyebrow lift, which in this case will more likely belong to an emotion of surprise or fear than other emotions such as sadness. Gathering the information of all the muscular movements can outline a forecast on the emotion.

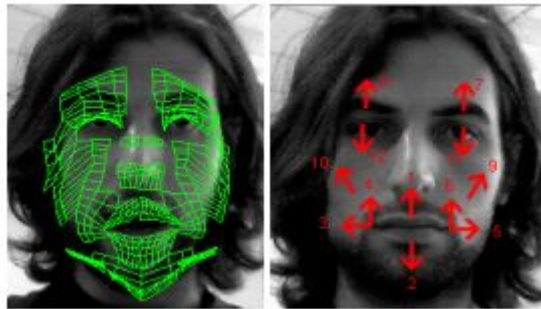


Figure B.3 Preview of algorithms for tracking facial expressions..

Use cases:

- Checking the effectiveness of promotional content. It is possible to diagnose the effectiveness of an advertisement or offer of products on the consumer by analyzing their facial response at the time of viewing. This will allow you to build an effective marketing strategy.
- Measurement of the degree of satisfaction against audiovisual events. Based on the response obtained by the viewers of an audiovisual content it is possible to obtain information on the degree of satisfaction of an event.
- Interactivity between player and game based on facial gestures. This provides a more level of interactivity on which to construct videojugabilistic experiences.
- Ease in human-machine interactions. Any device that requires an input of visual information, and specifically facial, of the user, may allow you to obtain information about your emotions so you can take certain actions based on it. This can be applied, for example to robotics.

CONCLUSION

-At the end of the project we must review those aspects that have been achieved and also those that need improvements or simply have not been achieved.

In the first place we aspired to build a tool that allows us to provide a new communication channel to those who suffer from a visual disability that prevents them from benefiting from visually manifested bodily communication. This has been achieved and makes available to the disabled user a new communication plan, which provides a method of interaction that has not been possible to date. This results in a possible new way of communicating since it will be possible to take into account fields

such as the appearance or facial expressions in the interactions carried out by the user, which will move the user to approach contexts and situations in which this is more important.

Secondly the passive or complementary feature to the user's daily actions has been achieved relatively. The tool is relatively invasive when it emits messages related to events or losses of the face of the interlocutor, this is mainly due to the facial location and expression algorithms. However, this aspect will improve considerably in scenarios that comply with solvency the environmental requirements referred to the light or position of the face. In addition, the updating of the algorithms of tracking and use of available resources is constant. New updates in this area are quite often available, which makes it possible to forecast an improvement in performance in this area in the medium and long term.

The third place corresponds to the accessibility that the application has to be acquired and used by blind users. This aspect is solved with relative solvency since, despite obtaining versions for the majority of the global mobile park, there is no version for iOS. This could be solved by acquiring the corresponding license as there is support for the mentioned operating system. In addition the technical resources used are not a barrier as these are not high in the current mobile phone market.

In the personal sphere this project has served me a lot in different fields.

I have known the technologies supported in the analysis of feelings and recognition of emotions in their different variants: voice, text and image. I now realize how much progress is being made in this field when I see mobile phones or marketing systems rely on them to automate market strategies or to improve security systems.

I have also learned about accessibility technologies available for all types of disabilities, not only visual but also hearing or cognitive disabilities. Gradually I've learnt to see more and more relationships between disabilities and technology, especially those of the last decades, that provides greater ease to learn concepts or to carry out interactions with the environment.

It is a very important field for people who suffer these ailments and to study and develop in the coming decades, with the importance that implies improving the quality of life of this group.

In short I have learned a lot about technology, disability, and also about technical skills in problems that I have had to solve in the development of the project. All this invites me to continue and to keep an eye on how the field of technologies for people with disabilities evolves.

BIBLIOGRAFÍA

- [1]"Instituto Nacional de Estadística", *Ine.es*, 2008. Disponible en:
<http://www.ine.es/jaxi/Datos.htm?path=/t15/p418/a2008/hogares/p01/modulo1/10/&file=01002.px> [Último acceso: septiembre 2017].
- [2]"Datos estadísticos — Web de la ONCE", *Once.es*, 2017. Disponible en:
<http://www.once.es/new/afiliacion/datos-estadisticos> [Último acceso: septiembre 2017].
- [3]*Salud ocular Salud ocular universal universal. Un plan de acción mundial para 2014-2019*. 2013. Disponible en: http://www.who.int/publications/list/universal_eye_health/es/ [Último acceso: septiembre 2017].
- [4]*Visual Impairment and Blindness 2010*. 2010. Disponible en:
http://www.who.int/blindness/data_maps/VIFACTSHEETGLODAT2010full.pdf [Último acceso: septiembre 2017].
- [5]"IDC: Smartphone OS Market Share", *www.idc.com*. Disponible en:
<http://www.idc.com/promo/smartphone-market-share/os> [Último acceso: septiembre 2017].
- [6]"Android Developers", *Developer.android.com*. Disponible en: <https://developer.android.com> [Último acceso: septiembre 2017].
- [7]C. Chetan R., *Text Based Emotion Recognition: A Survey*. Pune Institute of Computer Technology, Pune, Maharashtra, India: International Journal of Science and Research (IJSR), 2013.
- [8]K. Dai, H. J. Fell and J. MacAuslan, *RECOGNIZING EMOTION IN SPEECH USING NEURAL NETWORKS*. College of Computer and Information Science, Northeastern University, Boston, MA, USA: Speech Technology and Applied Research, Bedford, MA, USA.
- [9]M. El Ayadi, M. S. Kamel and F. Karray, *Survey on speech emotion recognition: Features, classification schemes, and databases*. 2011, pp. 572-587.
- [10]A. Azcarate, F. Hageloh, K. van de Sange and R. Valenti, *Automatic facial emotion recognition*. Amsterdam, Holanda: Universiteit van Amsterdam, 2005.
- [11]D. Matsumoto and H. Sung Hwang, "Reading facial expressions of emotion", <http://www.apa.org>, 2011. Disponible en:
<http://www.apa.org/science/about/psa/2011/05/facial-expressions.aspx> [Último acceso: septiembre 2017].
- [12]"Affectiva. Metrics.", *Affectiva.com*. Disponible en: <https://developer.affectiva.com/metrics/> [Último acceso: septiembre 2017].

- [13] *Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal*. 2011.
- [14] "Convenio de Munich sobre Concesión de Patentes Europeas, de 5 de octubre de 1973 (versión consolidada tras la entrada en vigor del Acta de revisión de 29 de noviembre de 2000).", *Oepm.es*, 2000. Disponible en:
http://www.oepm.es/cs/OEPMSite/contenidos/NORMATIVA/NormasSobrePatentes_MU_Topografias_CCP/NSPMTCCP_DerechoEuropeoPatentes/ConvenioMunichConcesionPatentesEuropeas_5_Oct_1973.htm [Último acceso: septiembre 2017].
- [15] "AccessWorld Magazine", *Afb.org*. Disponible en: <https://www.afb.org/> [Último acceso: septiembre 2017].
- [16] "FingerReader | Fluid Interfaces", *Fluid.media.mit.edu*. Disponible en:
<http://fluid.media.mit.edu/projects/fingerreader> [Último acceso: septiembre 2017].
- [17] L. KNFB Reader, "KNFB Reader", *Knfbreader.com*. Disponible en:
<http://www.knfbreader.com> [Último acceso: septiembre 2017].
- [18] "Announcing the KNFB Reader iPhone App", *YouTube*, 2014. Disponible en:
<https://www.youtube.com/watch?v=cS-i9rn9nao> [Último acceso: septiembre 2017].
- [19] "Color ID", *Play.google.com*, 2016. Disponible en:
https://play.google.com/store/apps/details?id=greengar.color.identifier&hl=es_419 [Último acceso: septiembre 2017].
- [19] "TapTapSee", *Play.google.com*, 2016. Disponible en:
<https://play.google.com/store/apps/details?id=com.msearcher.taptapsee.android&hl=es> [Último acceso: septiembre 2017].
- [20] "Be my eyes", *bemyeyes.com*, 2017. Disponible en: <http://bemyeyes.com> [Último acceso: septiembre 2017].
- [21] N. Holland, "An app to help blind people to 'see'", *BBC News*, 2017. Disponible en:
<http://www.bbc.com/news/av/magazine-39056979/an-app-to-help-blind-people-to-see> [Último acceso: septiembre 2017].
- [22] "Use TalkBack gestures - Android Accessibility Help", *Support.google.com*, 2017. Disponible en: <http://support.google.com> [Último acceso: septiembre 2017].